University of Alberta

Multivariate Statistical Monitoring of a High-Pressure LDPE and EVA Copolymer Industrial Process

by

Vikas Kumar ©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of **Master of Science**.

in

Chemical Engineering

Department of Chemical and Materials Engineering

Edmonton, Alberta
Fall 2002

National Library of Canada
Acquisitions and Bibliographic Services
395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale du Canada
Acquisitions et services bibliographiques
395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-81426-2

Canada

University of Alberta

Library Release Form

Name of Author: Vikas Kumar

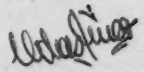
Title of Thesis: Multivariate Statistical Monitoring of a High-Pressure LDPE and EVA Copolymer Industrial Process

Degree: Master of Science

Year this Degree Granted: 2002

Permission is hereby granted to the University of Alberta Library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as hereinbefore provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.



Vikas Kumar
CME 536
University of Alberta
Edmonton, AB
Canada, T6G 2G6

Date: 26.07.2002

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled **Multivariate Statistical Monitoring of a High-Pressure LDPE and EVA Copolymer Industrial Process** submitted by Vikas Kumar in partial fulfillment of the requirements for the degree of **Master of Science** in *Chemical Engineering.*

Sirish L. Shah
(Dept. of Chemical & Materials Engineering)

U. Sundararaj
(Dept. of Chemical & Materials Engineering)

P. M. Hooper
(Dept. of Mathematical & Statistical Sciences)

Date: 15.07.2002

Abstract

The high-pressure LDPE industrial process operates under supercritical operating conditions, and therefore, it becomes imperative to devise schemes to prevent abnormal situations. Extreme deviations from normal operating region lead to conditions such as: loss of normal reaction, decompositions of the reactants and lost production due to the outages. Multivariate Statistical Process Control strategies operate on top of the DCS system to detect and diagnose abnormal process behavior in order to give the operators an opportunity to take preventive operational actions. Process engineers may also use this for off-line diagnosis of poorly understood processes for modifications in operating procedures.

In this work, data from a commercial LDPE/EVA copolymer high-pressure unit using an OPC server installed on the DCS is used to build empirical models and perform fault detection. Issues related to data transfer, preprocessing and filtering followed by development of PCA (Principal Components Analysis) models for various grades and first principles modelling of critical equipment will be presented. These models were used to detect process shifts. Process data from various real faults were considered and it was established that PCA could be employed to predict and diagnose process faults. The study closes with recommendations for online implementation of this process monitoring scheme and a new strategy based on multivariate statistics for product quality design.

I dedicate this thesis to Richard Bach and the spirit of his hero,
Jonathan Livingston Seagull

"Only those should be worshipped who are like ourselves but greater. The difference between the gods and me must be a difference only of degree."

- Swami Vivekanand

[15] The single word announced by all the Vedas Proclaimed by all ascetic practices, [The word] in search of which men practise chastity This word I tell [thee now] in brief. Om - this is it.

- Katha Upanishad Chapter II

Acknowledgements

My supervisors Prof.Sirish L Shah, Prof.U. Sundararaj and the technology researchers Larry Vande Griend and Dirk Hair from AT Plastics Inc. have had the greatest immediate influence on this thesis, and I am very grateful for all the hours I spent with all of them in discussions and consulting.

Thanks to the innovative brilliance and the patient listening skills of my supervisors that I could obtain a fair understanding of issues so critical to this work. I am also grateful to the teachers of the University of Alberta for the many skills they impart.

Thanks to the persistent follow up of Dirk Hair, Larry Vande Griend and Colin Sikora at AT Plastics that we could get the data server installed and running satisfactorily despite repetitive teething problems. Thanks are also due Carl, Dave, Emil, Brent and many other people at AT Plastics for their encouragement, humor and help in answering my incessant queries on the many intricacies of the process.

Thanks are due to my comrades-in-arms of the CPC group at the Department of Chemical and Materials Engineering, in particular Hari, Arun, Bhushan, Salim, Folake, Zhenghan and others for their good humor and abundant ideas. Many thanks to the computer wizards Bob and Jack for their help in a crucial program.

Through the magic of the satellite phone lines and the internet, several vital relationships have significantly helped to sustain me. Thank you to my family in India for all the good words and the unflinching support. Thank you also to best friends Sant, Vipan and Baljinder-I enjoy the talk on the philosophy of life with you.

I am blessed to have the family I do. Thanks Anita, Kanu and the latest in the family, Vikram, for loving me for who I am and supporting me for who I strive to become.

A very humble gratitude to my loving and kind spiritual Master Lord Krishna of Vrindavan for the inspiration and strength in my difficult and trying moments.

And of course, I am grateful for the scholarship I have received over the past two years from the University of Alberta, NSERC and AT Plastics Inc.

Contents

1 Introduction **1**

1.1 Multivariate Statistical Process Control 1

1.2 Organization of the Thesis 3

2 Basic Theory of Fault Detection and Diagnosis using Principal Components Analysis **5**

2.1 Introduction 5

2.2 Properties of PCA 5

2.3 Application 8

2.3.1 Monitoring charts based on Principal Components 9

2.3.2 SPE Plot 11

2.4 Fault Identification 12

2.5 Conclusions 13

3 High Pressure Polymerization Process **14**

3.1 Introduction and Scope of Work 14

3.2 Process Description 14

3.2.1 Reactor 15

3.2.2 Recycle and polymer processing 17

3.2.3 Special equipment 17

3.3 Data Server and Variable Selection 19

3.4 Special Events in Plant Operation 20

3.4.1 Cooler Cooking 20

3.4.2 Reactants Decomposition 21

3.4.3 Loss of Reaction 22

3.4.4 Grade Transitions 23
3.5 Concluding Remarks 23

4 Industrial Data Analysis **24**
4.1 Introduction 24
4.2 Data Preprocessing 25
4.3 Data Filtering 28
4.4 Secondary Compressor Model and Dynamic Flowrate Calculation . . 30
4.5 PCA Modelling 31
4.6 Conclusions 34

5 Fault Detection and Diagnosis: Decomposition **40**
5.1 Introduction 40
5.2 Model Development 42
5.3 Detection and Root Cause Diagnosis of Event 47
5.4 Conclusions 54

6 Fault detection and Diagnosis: Loss of Reaction **55**
6.1 Introduction 55
6.2 Model Development 57
6.3 Detection and Root Cause Diagnosis of Event 59
6.4 Conclusions 69

7 Product Quality Design **70**
7.1 Introduction 70
7.2 Multivariate Statistical Tools for Product Quality Design 71
7.3 Correlation studies 73
7.4 PCA model: Information on dominant factors 79
7.5 Conclusions 79

8 Conclusions **81**
8.1 Contributions of this Thesis 81
8.2 Recommendations for Future Work 82

Bibliography 84

A Process Variables Logged by Data Server 89

B Raw Data File 91

C Cooler Cook Model Loadings 93

D Loss of Reaction Model Variables and Loadings 95

E Correlation Matrix for 18% VA Product Cooler Cook 97

List of Figures

2.1 Projection of the Data-Matrix 7
2.2 Score plot of PC-1 vs PC-3 showing the control ellipse and the dense scatter of 'NOC' operation 9
2.3 Hotelling T^2 Plot 10
2.4 SPE Plot 11

3.1 Autoclave Reactor 15
3.2 Process Flow Diagram 16
3.3 Schematic showing the placement of the Data Server on the existing DCS 19

4.1 Raw Data Plot 30
4.2 Filtered Data Plot 31
4.3 Secondary Compressor Model for Flowrate Calculation 35
4.4 Eigenvalue Scree Plot 36
4.5 Cumulative Variance Plot: Seven PC's 36
4.6 Loadings Plot:PC 1 38
4.7 Loadings Plot:PC 2 38
4.8 Correlation Matrix 39

5.1 Cooler Cook comparison with normal operation 43
5.2 Cooler Cook comparison with normal operation trends 44
5.3 Plot showing abnormal variation of High Pressure Separator pressure affecting model threshold 45
5.4 Plot showing abnormal variation of Product Cooler outlet temperature affecting model threshold 45

5.5 SPE of Decomp Model showing appended data from six cooler cooks 46
5.6 Decomp Model Scores 46
5.7 Decomp Model Eigenvalue Scree Plot 48
5.8 Decomp Model Cumulative Variance captured by PC's 48
5.9 SPE Plot showing violation of threshold during seventh cooler cook . 49
5.10 (a)Decomposition: Scores plot showing the violation of the threshold ellipse. (b) The lower figure is a zoom-in of the area where the violation occurs 49
5.11 Zoom-in of the area where violation of the SPE threshold occurs and the point of decomposition 50
5.12 SPE Contribution Plot:Bars show the %contribution of each variable towards the violation 50
5.13 FLOWRATE in kg/hr versus sampling time for the six earlier cooler cooks and for the final cooler cook which ended in a decomposition . 52
5.14 Trend Plot of percentage contributions of all 35 variables during the violation of the SPE threshold.It is hard to distinguish the top five contributors from such a combined plot. 53
5.15 Trend Plot of the top five contributors to the violation of the SPE threshold. 53

6.1 Trend Plot of Catalyst Pump Outputs in Zone1B showing Pump Changeover 57
6.2 Trend Plot of Temperature Loops of Reactor 58
6.3 Scree Plot:Five PC's selected for Loss of Reaction model 59
6.4 Plot showing cumulative variance captures by Five PC's selected for Loss of Reaction model 60
6.5 SPE Plot of Model 61
6.6 PC5 violation of threshold:Loss of Reaction 62
6.7 Score Plot showing violation of threshold:Loss of Reaction 62
6.8 SPE contribution plot sample instant 9619 showing the flowrate and compressor variables as the top contributors:Loss of Reaction 63
6.9 Hotelling's Plot showing violation of threshold:Loss of Reaction . . . 64

6.10 SPE:Loss of Reaction owing to catalyst pump failure 64
6.11 SPE:Loss of Reaction owing to catalyst pump failure 65
6.12 SPE contribution plot sample instant 10075 at which sharp rise starts in SPE starts:Loss of Reaction 66
6.13 SPE contribution plot matching with sample instant at which Hotelling's threshold is also violated:Loss of Reaction 66
6.14 Trend Plot of Contribution of all variables in model:Loss of Reaction at monitoring point 121 67
6.15 Trend Plot of interstage pressure PI51176.PV :Loss of Reaction . . . 68
6.16 Trend Plot of Flowrate:Loss of Reaction 68

7.1 A Bar chart for MI correlation with plant parameters 74
7.2 The organizational chart for correlation indices of MI with plant parameters. 75
7.3 Correlation color matrix: Notice the correlations of parameters with MI (variable no.2), the highly correlated compressor variables (nos.15 to 19), the highly correlated reactor temperatures (nos.20 to 24), the highly correlated variables of Zone1 (nos.26 to 29). 78

A.1 Process Variables logged by server 90

B.1 Raw Data File Format 92

C.1 Cooler Cook Model Loadings 94

D.1 Loss of Reaction model variables and loadings 96

List of Tables

4.1 Loadings and Tag Configuration of the model with seven PC's 37

5.1 Decomp Model PC's and their significance 47

6.1 Table showing PC#,associated eigenvalue, variance captured,variables covered by the various PC's 61

Chapter 1

Introduction

1.1 Multivariate Statistical Process Control

Polymer manufacturing has seen tremendous growth in the last half century with scores of products or grades being manufactured for specific applications from single process units. This requirement of multi-grade production has led to research in industry to search for mechanistic or empirical models for a greater insight into complicated processes. Mechanistic models based on reaction kinetic analysis have been reliable as far as process and approximate product design is concerned, but fall short of exactly defining the correlations under various process operating conditions required for manufacturing multiple grades. As a result, technical research and development groups have developed recipe based production wherein operating conditions are defined for grades which are then fine tuned with operating experience. The drawback of this approach is that new grade development requires costly trials and in case of operation breakdown or faults, there are very few tools available to diagnose the reasons for the situation leading to the unexplained outage. At most, historical univariate plots are obtained and the best possible explanation based on reaction kinetics is made. It is fairly well understood by research and operation groups that many of the process variables in a unit are highly correlated and hence process monitoring should be based on techniques which look at all the critical variables simultaneously. This is only possible by using multivariate statistical tools for monitoring and product design.

In the modern day process industry, hundreds of process variables are being sampled and stored for every unit. Whether the control strategy be implemented via

Distributed Control Systems (DCS) or multiple-loop Programmable logic controllers (PLC), these systems are capable of generating process variable data sets which are overwhelming in size. Most industrial chemical processes are multivariate in nature and multivariate statistical techniques are being employed to condense these data sets. The condensed data is subsequently used for knowledge extraction to improve process monitoring and closed loop control. We can take advantage of the fact that many of the process variables are correlated and therefore, it may be be possible to develop relationships between them. Additionally, since most of the data being recorded is redundant, informed decision-making is done based on process knowledge and by employing multivariate statistical techniques to select only useful information for further analysis.

The most common methods for data reduction are Principal Components Analysis(PCA) and Partial Least Squares(PLS). PCA is used for large data sets of processes containing significant redundancies. It compresses the data set to fewer pseudo variables or scores which capture most of the variance of the original data matrix. PLS also known as projection to latent structures, is a dimensionality reduction technique. PLS uses two blocks of data and systematically compresses them. It maximizes the covariance between the predictor variable matrix $\mathbf{X}$, comprising process variables and the predicted matrix $\mathbf{Y}$, comprising quality variables.

PCA is a common procedure to reduce the dimensionality of the process variable space and was discovered by Pearson (1901), and further developed by Hotelling (1933, 1947), followed by recent researchers (Wold and Geladi 1987, Mardia and Bibby 1982, Kresta and MacGregor 1991, Jackson 1991). Application of PCA has been demonstrated by both academic and industrial researchers. While substantial work has been done and published on data collected from simulations (Himes and Georgakis 1994, Kasper and Ray 1992, Ku and Georgakis 1995, MacGregor 1994, Raich and Cinar 1994, Raich and Cinar 1995, Luo and Himmelblau 1999), detailed applications of the techniques on industrial data have not been published as widely, probably due to the proprietary nature of the data. One of the earliest works on PCA applications on industrial data was published by researchers from a major manufacturer and marketer of chemical products Showa Denko K. K., of Japan (Moteki and Arai 1988*a*). This work was oriented towards improving the understanding of

the behavior of a high pressure polymerization process for product quality design. Journal and conference proceedings articles have also been published by researchers from industry , notably DuPont, and academics on how PCA can be applied to industrial data for process monitoring(MacGregor and Kiparissides 1991, Piovoso and Pearson 1992, Kosanovich and Nomikos 1994, Piovoso and Kosanovich 1994, Wise and Gallagher 1996, Kourti and MacGregor 1998), but none of these give process details and special operating events of the plant being studied. From the collection of published work on applications of PCA and PLS for process monitoring, fault detection and diagnosis over the last one and a half decades, few industrial applications have been published using real process data. The exceptions are Skagerberg 1992 and Zheng 2001, who have also described operation event history which have lead to faults. (Skagerberg and Jan 1992, Zheng and Chen 2001).

In the current work, PCA is used to model a high pressure polymerization process, using data from a server on the DCS of an operating plant, for use in fault detection, diagnosis and control. Correlation studies are carried out to aid product quality design. Plots such as the Squared Prediction Error(SPE) and Hotelling T^2, which condense most of the process variation into single plots, are used to monitor the process. This study mainly employs multivariate statistics to build empirical models; and mechanistic modelling is used to derive the mass flowrate of reactants through the reactor.

1.2 Organization of the Thesis

The thesis is concerned with issues of PCA modelling, fault detection and diagnosis for a polymerization process. Each chapter begins with an introduction giving the reader an overall picture of the subject present therein.

Chapter 2 lays down the theoretical background to PCA, its properties and the charts used for monitoring the polymerization process. SPE plots utilized as a single window for observing the process are presented as an effective tool for fault detection and subsequent root cause diagnosis.

This research work on high pressure polymerization fault detection and diagnosis has been supported by AT Plastics Inc., Edmonton. *Chapter* 3 gives a process de-

scription of the high pressure *ICI* process,used by AT Plastics to manufacture many different grades of *LDPE*(Low density polyethylene polymer) and *EVA*(ethylene vinyl acetate) copolymers. The objectives of this research study are defined in this chapter. Various special operating procedures are also described in some detail; these procedures will be referred to in later chapters.

Chapter 4 details the standard procedures used for data pre-processing. These procedures are applied uniformly to data from all grades before performing PCA modelling. This chapter also describes the issues involved in doing first principles modelling of a hyper compressor handling fluid at supercritical state. The modelling was done to derive mass flow rate through the compressor, and hence, the reactor.

Chapter 5 is concerned with a highly researched and poorly understood phenomena of reactants decomposition in the high pressure polymer industry; the occurrence of which generally causes a unit outage. The data used for the study is from a real decomposition in the plant that occurred while cooler cook operation taking place. Issues of model development using data from similar grades and fault prediction are also described here.

Chapter 6 discusses another frequent cause of unit shutdown in the polymerization industry which is loss of reaction. The case in which a catalyst or initiator pump failure leads to a loss of reaction is discussed. Issues of model development using data from similar grades and fault prediction are described.

Chapter 7 presents correlation charts which can be used for development of simple monitoring techniques and product quality design.

Chapter 8 gives the conclusions and future work. Within the chapter are suggestions for improvement in process control at AT Plastics Inc. and a strategy for implementation of an on-line fault detection and diagnosis scheme using Multivariate Statistical Process Control.

Chapter 2

Basic Theory of Fault Detection and Diagnosis using Principal Components Analysis

2.1 Introduction

The modern process industry is heavily instrumented and has large historical databases of process variables. The problem is to find effective ways of extracting useful information from this data. The basic idea behind PCA is data compression and knowledge extraction. This is done by finding fewer factors than the original data set which will describe the major trends in the process thereby reducing the number of variables to be monitored. Since the information in a process does not lie in a single process variable, PCA is used to capture the variance information in process variables with respect to each other or in other words how they 'co-vary' (Wise and Gallagher 1996). In such a condition, information is extracted from a data set to reduce the number of variables to be monitored. Using this dimension reduction technique, models are set up using historical data. Special events are detected when the process variables deviate significantly from the 'in control' model developed using historical data.

2.2 Properties of PCA

The objective is to model a *single* block of data, $\mathbf{X} \epsilon \Re^{nxm}$ with n observations and m variables. The main idea is to explain the covariance structure through a *few linear combinations* of the original variables. We take advantage of the fact that the original

m variables are correlated and most of the variability can be explained by as few as q pseudo or latent variables ($q \ll m$), also called q principal components. Consider an observation at a single time instant x with m variables. The m linear combinations of these variables are also called scores and are defined as following,

$$t_1 = p_{11}x_1 + p_{21}x_2 + ... + p_{m1}x_m \tag{2.1}$$
$$t_2 = p_{12}x_1 + p_{22}x_2 + ... + p_{m2}x_m$$

$$t_m = p_{1m}x_1 + p_{2m}x_2 + ... + p_{mm}x_m$$

where $x_1, x_2, ..., x_m$ are the m observations at a sampling instant from the data matrix $\mathbf{X}$. In terms of capturing the variance, the first principal component of $\mathbf{x}$ is that linear combination that has maximum variance and is defined as,

$$t_1 = \mathbf{x}\mathbf{p_1} \tag{2.2}$$

and is subject to $\mid \mathbf{p_1} \mid = 1$ The second principal component is that linear combination which has the next greatest variance,

$$t_2 = \mathbf{x}\mathbf{p_2} \tag{2.3}$$

and is subject to $\mid \mathbf{p_2} \mid = 1$. The vectors p_1 and p_2 are orthogonal to each other and hence, both t_1 and t_2 are uncorrelated. On considering n samples of observations, with m variables, forming a data matrix $\mathbf{X}$, which is a subset of the population data matrix and has been mean autoscaled, we may find a sample correlation matrix $\hat{\mathbf{\Sigma}}$ from which the corresponding eigenvalues λ_i, and eigenvectors $\mathbf{p_i}$ with $i = 1, 2, ..., m$ are evaluated by solving an eigenvalue decomposition of the correlation matrix.

$$\hat{\mathbf{\Sigma}} = \mathbf{S} = (n-1)^{-1}\mathbf{X^T X} = \mathbf{P}\Lambda\mathbf{P^T} \tag{2.4}$$

The principal components loading vectors are the orthonormal column vectors of the matrix $\mathbf{P} \in \Re^{m \times m}$ from the above solution and $\Lambda \in \Re^{m \times m}$ contains the non-negative real eigenvalues of decreasing magnitude ($\lambda_1 \geq \lambda_2 \geq ...\lambda_m \geq 0$) along

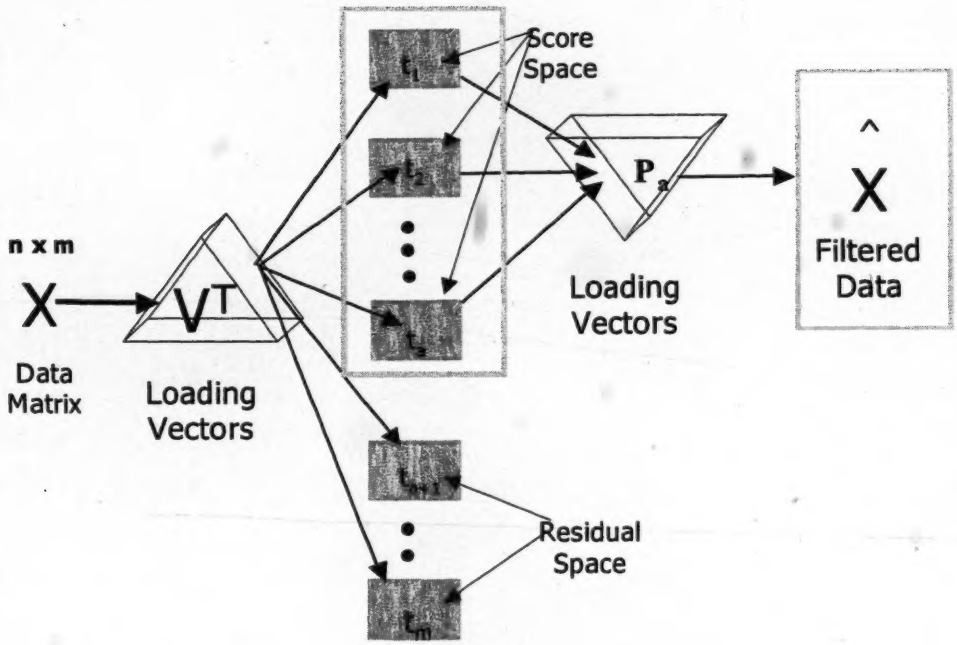


Figure 2.1: Projection of the Data-Matrix

its main diagonal with zero off-diagonal elements. This is equivalent to performing singular value decomposition(SVD) on the data matrix $\mathbf{X}$ (Chiang and Braatz 2000). It should be noted that the eigenvalues are the variances of the principal components (PC's) or in other words,

$$Var(\mathbf{t_i}) = \lambda_i \tag{2.5}$$

In effect PCA decomposes the observation matrix X as,

$$\hat{\mathbf{X}} = \mathbf{T}\ \mathbf{P^T} = \sum_{i=1}^{q} \mathbf{t_i p_i^T} \tag{2.6}$$

where $\hat{\mathbf{X}}$, is the approximation of the original data matrix $\mathbf{X}$. PCA is scale dependent and the matrix is auto-scaled, or by another appropriate way depending upon the variables in the data matrix as explained in §4.2.

Since this is a dimension reduction technique we choose only a certain number of principal components $q \ll m$, which explain most of the variability in the data. Information on the number of PC's to be used has been given by Wold 1978, Geladi and Kowalski 1986 and Jackson 1991. A useful method for selection of those number

of PC's which minimizes the loss function is via the following expression:

$$loss function = \sum_{i=1}^{m} \frac{SPE_i}{q-1} \tag{2.7}$$

where SPE is calculated as,

$$SPE_x = \sum_{i=1}^{m} (\mathbf{x_{i,new}} - \mathbf{\hat{x}_{i,new}})^2 \tag{2.8}$$

This loss function would keep going down with respect to the value of q. The loss function is plotted against various values of q to locate that value of q where the slope starts turning and is no longer linear.

In this work, the preliminary selection of the number of PC's in the model is done by performing a scree test of the eigenvalues, λ_i of the covariance matrix. The point where the profile is no longer linear is chosen as the cut off point for the number of PC's. In addition, the number of PC's in the model should explain a minimum of 80% of the variance in the data. In some cases cross validation with another 'normal' data set is done for the final selection of the number of PC's adopted in the model.

The following properties hold for PCA,

1. $var(\mathbf{t_1}) \geq var(\mathbf{t_2}) \geq .... \geq var(\mathbf{t_m})$
2. $mean(\mathbf{t_i}) = 0; \forall i$
3. $\mathbf{t_i^T t_k} = 0; \forall i \neq k$
4. **No other orthogonal expansion of q components exists which captures a higher cumulative variance of the data set.**

2.3 Application

After a PCA model is built based on Normal Operating Conditions (NOC) when only common cause variation exists, we may predict the future behavior of the plant by comparing the present operation with this normal plant model. The tools used to achieve the above are score plots, Hotelling T^2 plot and the squared prediction error or the SPE plot.

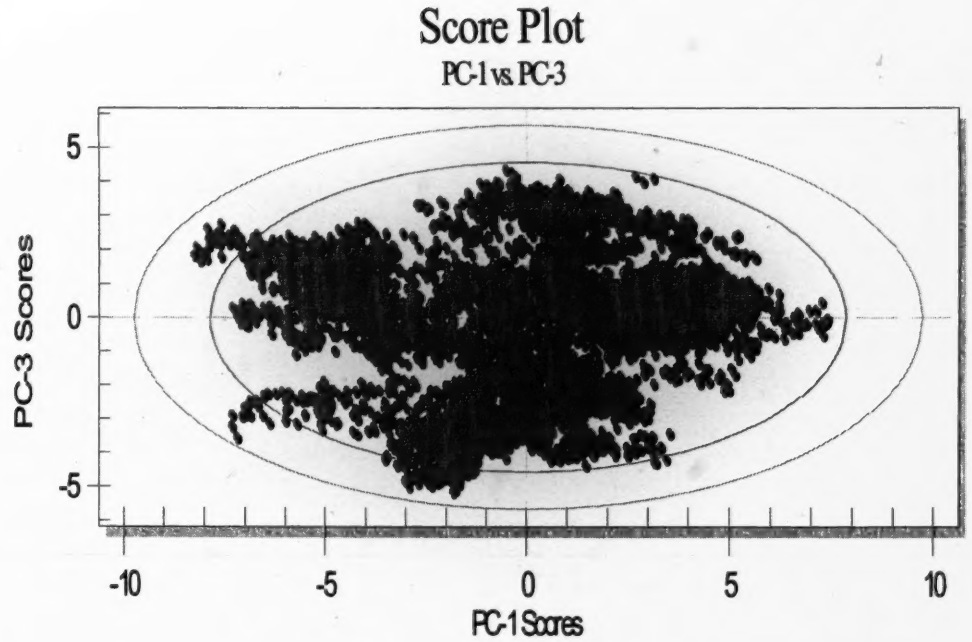


Figure 2.2: Score plot of PC-1 vs PC-3 showing the control ellipse and the dense scatter of 'NOC' operation

2.3.1 Monitoring charts based on Principal Components

The steps for developing the aforesaid plots are as follows:

New observations are projected using the model to get new scores,

$$\mathbf{t_{i,new}} = \mathbf{x_{new}} \mathbf{p_i} \tag{2.9}$$

A **score plot** is then obtained by plotting any two principal components against each other, e.g., $\mathbf{t_1}$ *vs* $\mathbf{t_3}$ (or PC-1 vs PC-3). See Fig.2.2 for an 'in-control' score plot. During normal operation 99% of the plotted points in a score plot will lie within a threshold contour called the control ellipse which is calculated by considering the bivariate distribution of the two principal components scores. This threshold is a solution to the Hotellings T^2 generalization of the squared distance for a certain level of significance.

The **residuals** are calculated as,

$$\mathbf{e_{new}} = \mathbf{x_{i,new}} - \hat{\mathbf{x}}_{\mathbf{i,new}} = \mathbf{x_{new}} - \mathbf{t_{i,new}} \mathbf{p_i}^{\mathbf{T}} \tag{2.10}$$

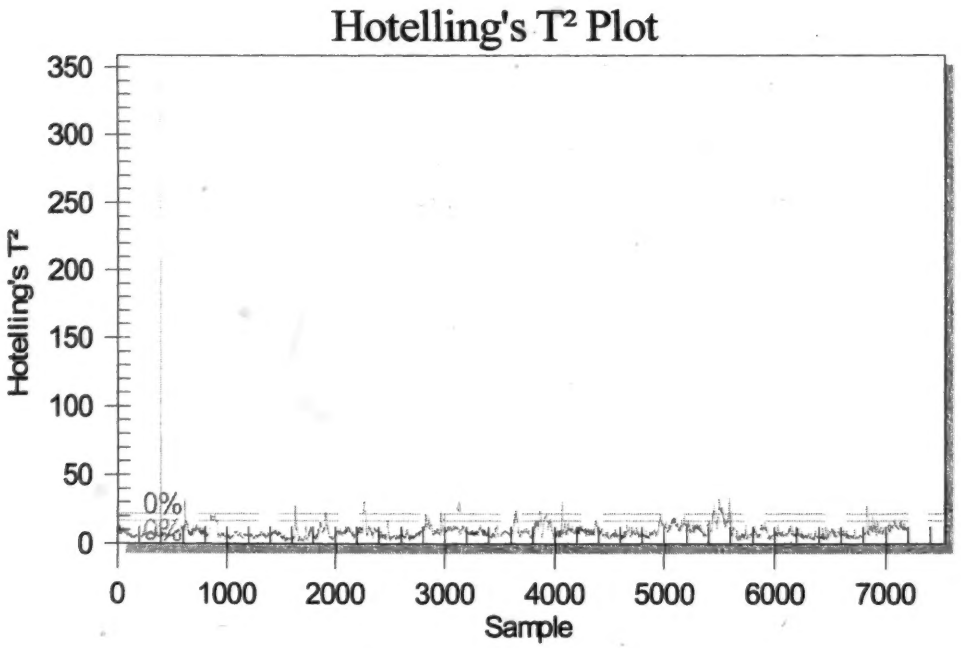


Figure 2.3: Hotelling T^2 Plot

Thereafter the new projected scores and the residuals are used to calculate statistics for monitoring the process.

The **Hotelling** T^2 statistic is calculated as follows:

$$\mathbf{T^2} = \sum_{i=1}^{q} \frac{\mathbf{t_i^2}}{\mathbf{s_{t_i}^2}} = \sum_{i=1}^{q} \frac{\mathbf{t_i^2}}{\lambda_i} \tag{2.11}$$

where q is the number of PC's in the model. We scale each $\mathbf{t_i^2}$ with the reciprocal of its variance to ensure that each PC term plays an equal role in the calculation of $\mathbf{T^2}$ statistic irrespective of the amount of variance it explains in the $\mathbf{X}$ data matrix (See Fig.2.3).

The PC's from $(q+1, ..., m)$ explain only a small amount of the total variance in the data matrix $\mathbf{X}$ and generally represents random noise. Since the variances of the lower PC's are small, scaling the scores with these small variances may blow up the $\mathbf{T^2}$ statistic and indicate an out of control situation. Hence the $\mathbf{T^2}$ based statistical test on the first q PC's provides a more robust test for detecting deviations in the process.

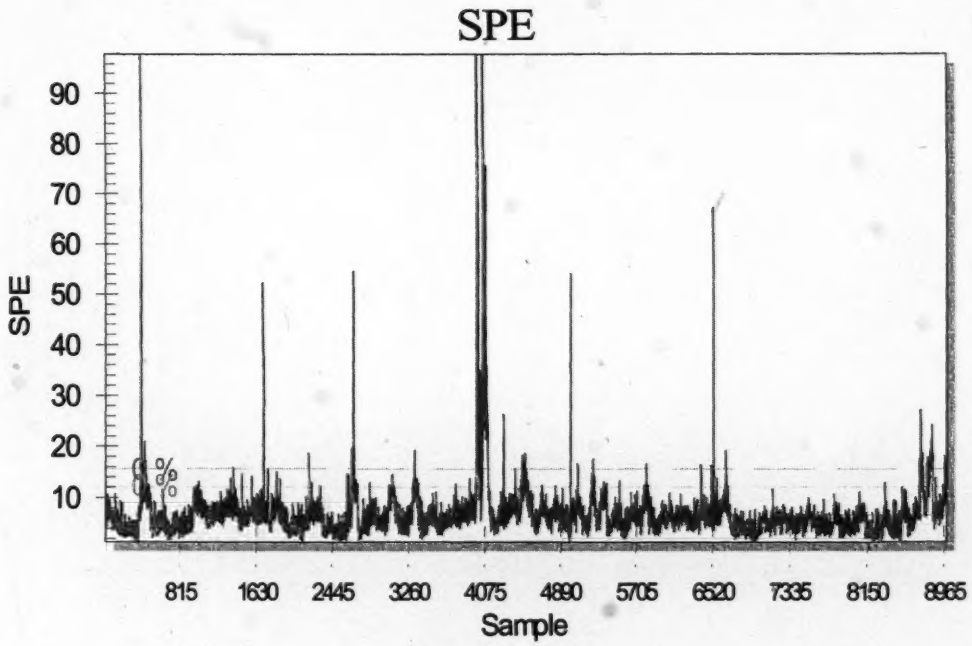


Figure 2.4: SPE Plot

2.3.2 SPE Plot

Abnormal events may also be detected by computing the Q-statistic or the Squared Prediction Error (SPE_x) which is the sum of the squared residuals of the new observation vector at every sampling instant. The procedure for calculating the SPE_x plot has been explained by (Jackson 1991) and (Kresta and MacGregor 1991). The control limits for the SPE_x plot are computed from historical data, using approximate results for the distribution of the quadratic forms (Jackson 1991). When an abnormal event occurs in the process plant, it results in a change in the correlation structure of the data matrix $\mathbf{X}$ and the event will be detected by a high value of SPE_x. This means that the projection model is no longer valid for that observation. The SPE_x is calculated as follows,

$$SPE_x = \sum_{i=1}^{m} (\mathbf{x_{i,new}} - \mathbf{\hat{x}_{i,new}})^2 \tag{2.12}$$

The value of SPE_x during the 'in control' state is small and below the control limits (See Fig.2.4). Changes in SPE_x in this condition represents fluctuations that are due

to common cause variations. Another definition for the SPE_x is the total sum of variations in the residual space. Work on the threshold limit calculation for SPE_x was done by Jackson and Mudholkar 1991, and is calculated as follows,

$$Q_\alpha = \theta_1 [\frac{h_0 c_\alpha \sqrt{2\theta_2}}{\theta_1} + 1 + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1^2}]^{\frac{1}{h_0}} \tag{2.13}$$

where $\theta_i = \sum_{j=a+1}^{m} \lambda_j^i$, for $i = 1, 2, 3$, and $h_0 = 1 - \frac{2\theta_1\theta_3}{3\theta_2^2}$, with c_α being the normal deviate corresponding to $(1 - \alpha)$ percentile in a normal distribution.

2.4 Fault Identification

In the event of detecting a violation of the $\mathbf{T^2}$ threshold, the next step is to calculate the contribution of each process variable to the individual scores at the n^{th} observation step. The following procedure is followed, (MacGregor and Kourti 1995, Miller and Heckler 1998):

1. The normalized scores, $\frac{\mathbf{t_i^2}}{\lambda_i}$ are calculated.

2. Identification of those $'k'$ scores which are causing the process to move out of the 'normal-operating-condition'(NOC) is done by finding the scores with $\frac{\mathbf{t_i^2}}{\lambda_i} > \frac{1}{\mathbf{q}}(\mathbf{T_\alpha^2})$

3. The contribution of each process variable to the 'out-of-NOC' is calculated as follows,

$$Cont_{i,j} = \frac{\mathbf{t_i}}{\lambda_i} \mathbf{p_{i,j}} [\mathbf{x_j} - \mu_j] \tag{2.14}$$

   where $\mathbf{p_{i,j}}$ is the $(\mathbf{i}, \mathbf{j})^{th}$ element from the loading matrix comprised by the subset of 'k' PCs. The negative contributions are neglected.

4. The total contribution of the j^{th} process variable, x_j, is evaluated.

$$Cont_j = \sum_{i=1}^{k} cont_{i,j} \tag{2.15}$$

   where $i = 1, 2.., k$ since we are identifying those $'k'$ scores which are causing the process to move out of of 'NOC' region(refer to step 2 above).

5. The contributions of all process variables are then plotted on a single graph.

An alternative scheme for fault identification is based on quantification of the total variation of each of the process variables in the residual space (Wise at al,1989). This method is based on an observation vector at each sampling instant and is the normalized error vector,

$$Res_j = \frac{\mathbf{r_j}}{\hat{\mathbf{s}}_\mathbf{j}} \tag{2.16}$$

where r_j is the jth variable of the residual vector and $\hat{s}_j$ is the standard deviation of the jth variable in the training set of observations (Chiang and Braatz 2000). The normalized error vector is then plotted on a single graph in the form of percentage contribution to the SPE_x. The highest contributors are the process variables causing the process to go out of 'NOC' (Kresta and MacGregor 1991, Himes and Georgakis 1994, MacGregor 1994).

2.5 Conclusions

Projection methods are very effective in condensing, analyzing, monitoring and diagnosing operational problems. In this thesis, the statistical technique applied for fault detection and diagnosis is PCA. This technique is a powerful predictor of slow or incipient faults which are frequent in many process units. The main tools used are the Hotelling T^2 statistic, the $2D$ score plot and the SPE plot.

Chapter 3

High Pressure Polymerization Process

3.1 Introduction and Scope of Work

This project on multivariate statistical process monitoring, fault detection and diagnosis of LDPE and EVA production, is based on data obtained from reactor R-V (read R-five, the fifth reactor) unit of AT Plastics Inc., Edmonton, Canada. The company is a pioneer in the production of special resins and has been in operation for over five decades. R-V is the most recently built unit of five manufacturing units at the Edmonton site and is based on the High Pressure ICI Autoclave process for manufacturing LDPE and EVA copolymers. The unit produces more than 35 grades of specialty EVA resins and is controlled using a TDC3000 Honeywell DCS. The process description and special operating procedures relevant to the current work are explained in the following sections.

3.2 Process Description

Autoclave Process: This process, popularly known as the ICI process, is characterized by supercritical operating conditions with pressures in the range of $150 - 300$ MPa and relatively moderate temperatures of $200 - 300°C$. The polymerization is carried out in stirred autoclave reactors with negligible heat loss through the walls of the reactor. The same unit may also be used for co-polymerization where the comonomer is injected at the final compressor suction. The unit has successive trains of compressors with inter-cooling and vapor-liquid separation equipment before the

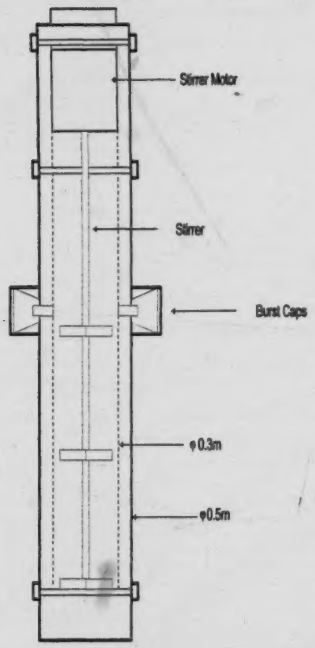


Figure 3.1: Autoclave Reactor

feed enters the reactor in four separate streams. The unconverted gases from the reactor are cooled and recycled after separating the liquid in vapor-liquid separation equipment which are cyclonic or pressure let down systems. Since the conversion of ethylene to polymer is determined by the temperature difference across the reactor, the feed gas is brought in at low temperatures and the reaction temperatures are set fairly close to the temperature where ethylene decomposes to carbon, hydrogen and methane. If decomposition occurs, it is strongly exothermic and on initiation will proceed rapidly to increase pressure beyond the vessel design pressures and have to be relieved through special relief systems. The decomposition products are diluted with water before they are vented to atmosphere in order to avoid aerial decomposition.

3.2.1 Reactor

A rotary stirrer driven by an induction motor installed internally in a high-pressure tubular reactor is typical of reactors used in the ICI process (See Figure 3.1). The polymerization is carried out by a free radical mechanism initiated with a mixture of

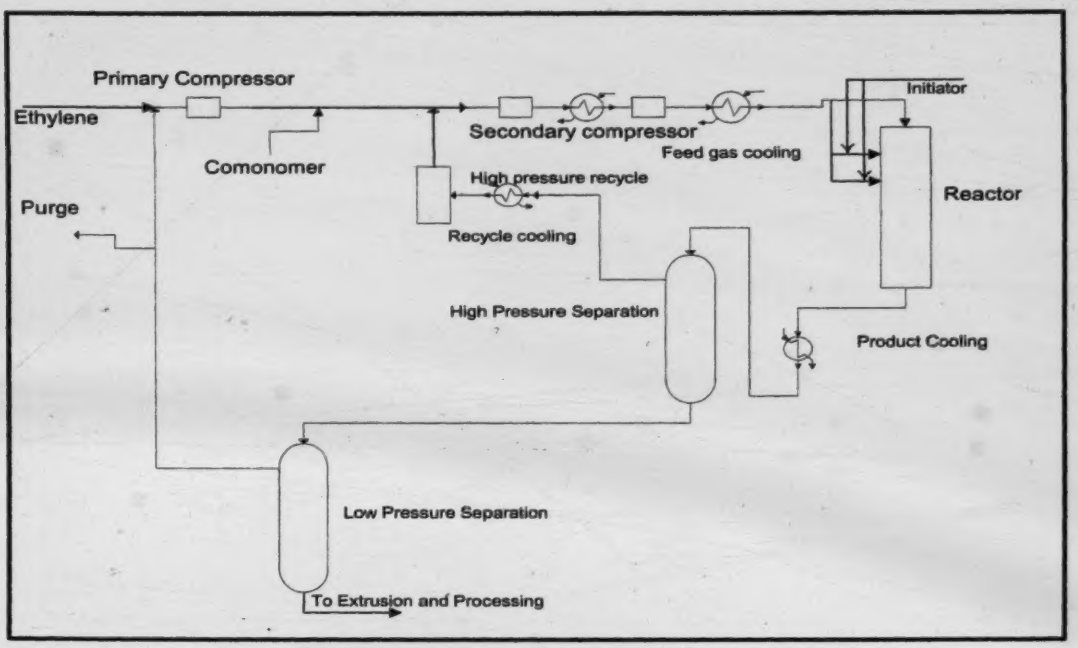


Figure 3.2: Process Flow Diagram

organic peroxides introduced in specific zones of the reactor or into the feed gas. To control reaction temperature and to avoid hot spots caused by non-uniform mixing, it becomes imperative to have excellent mixing through out the reactor. The motor, its housing and the stirrer have to be designed to avoid local hot spots which may cause ethylene decompositions. The stirrer speed and mechanical design are chosen to avoid polymer build up on the shaft and to achieve high turbulence. Polymer build up on the stirrer shaft leads to quality problems, shaft vibration and potential hot spots. The stirrer has a proprietary design of multiple paddles and baffles arrangement. The baffles also act as partitions between reaction zones in the reactor. Initiator is introduced in each zone and each zone has independent temperature control loops. The residence time varies from 10 seconds to over a minute. Single pass conversion varies from anywhere from 9 to 22 % (See Fig 3.2). The reactor is jacketed and circulating hot oil is used for start-up and initiation of the polymerization reaction. The reaction is then sustained by the exothermicity of the reaction and initiator addition. In case of decomposition, multiple bursting discs on the reactor vessel wall

ensure the release of decomposition products from the reactor. The reactor is not the only equipment in which decomposition occurs. Local hot spots in the secondary compressor or in other high-pressure zones may also lead to decompositions. Aside from hot spots, small amounts of oxygen may also initiate decompositions; hence it is essential to seal the low-pressure section and purge thoroughly after an annual shut-down or any outage where the system has been depressurized. **Decomposition is a critical fault, and industry is looking for ways to predict it so that it can be avoided or its effects minimized.**

3.2.2 Recycle and polymer processing

The reactor effluent is cooled and sent to a high-pressure separator to separate the polymer from the unreacted feed gases. The overhead stream from the separator comprising unreacted feed is recycled through a series of coolers and cyclonic separators to the secondary compressor suction. Cyclonic separation is carried out to remove low molecular weight polymer from the recycle gases. This stream combines with fresh feed at the secondary compressor suction. The polymer from the high-pressure separator is sent to a low-pressure separator (See Fig.3.2). Then the separated polymer is sent to the extruder. The low-pressure off-gas is partly vented to control impurities accumulation in the circulating gas environment and the balance is recycled to primary compressor suction. Vinyl acetate copolymerizes with ethylene and is added at secondary compressor suction. The recipe may have up to 40wt% vinyl acetate, and as the grade is changed, there will be a change in plant operating conditions.

3.2.3 Special equipment

Initiator Pumps: Initiators for this process are generally mixtures of peroxides in solvent that are injected into the reactor using multi-cylinder, mechanically driven pumps. The multi-cylinder pumps are usually double acting and driven hydraulically with stroke speed reset by the temperature in the reactor. This control loop is a difficult control problem and an interruption or unsteady flow of catalyst may lead to decomposition or a loss of reaction. A loss of reaction occurs when the polymerization slows down and the monomers are not converted. The major problem in these pumps is leaking seals. This occurs because the solvents used with the the peroxides attack

the sealing material (which is usually made of PTFE). Some peroxides may also deposit on the packing or the internal valves of the pump. Excessive deposit may damage the seals and may also cause the ball check valve to get stuck leading to a loss of reaction.

High pressure reciprocating compressors: In an LDPE plant a primary compressor is usually a two stage machine and raises the pressure of the feed gases to $15 - 30$MPa. A secondary compressor raises the pressure to $150 - 300$MPa. The variation of density of ethylene and its mixtures with vinyl acetate leads to different compression ratios over the two stages of the secondary compressor. This results in large pressure differences over the compressor. The major problems to be considered are leakages from the compressor seals, and the cyclical fatigue failures of the cylinders and valves. Small amounts of additional low molecular weight polymer is formed and deposits in the piping, equipment and compressor internals. This can to lead to nucleation raising the probability of forming cavities supporting generation of hot spots where ethylene break down may occur.

In the high pressure process, there are two compressors in series which raise the pressure of the feed gas to supercritical operating pressures. These are the primary compressor and the secondary compressor. The primary compressor raises the pressure of primarily fresh feed gas with a small component of recycle gas (about $6Wt.\%$ of the total) to $15 - 30$MPa. The secondary compressor has higher volumetric capacity compared to the primary compressor since it compresses both the fresh feed gas from the primary compressor and recycle gas which is around $74Wt.\%$ of the total throughput through the reactor. The outlet pressures achieved are supercritical and are in the range of $150 - 300$MPa. In addition to the mechanical challenges of design and maintenance, there are some major problems associated with the compression process. These are:

1. Low molecular weight polymer is formed at these conditions which deposits in the associated piping and equipment.

2. Break down of ethylene to carbon, hydrogen and methane might occur if there are local hot spots, i.e. decomposition. This leads to equipment damage or activation of pressure relief devices.

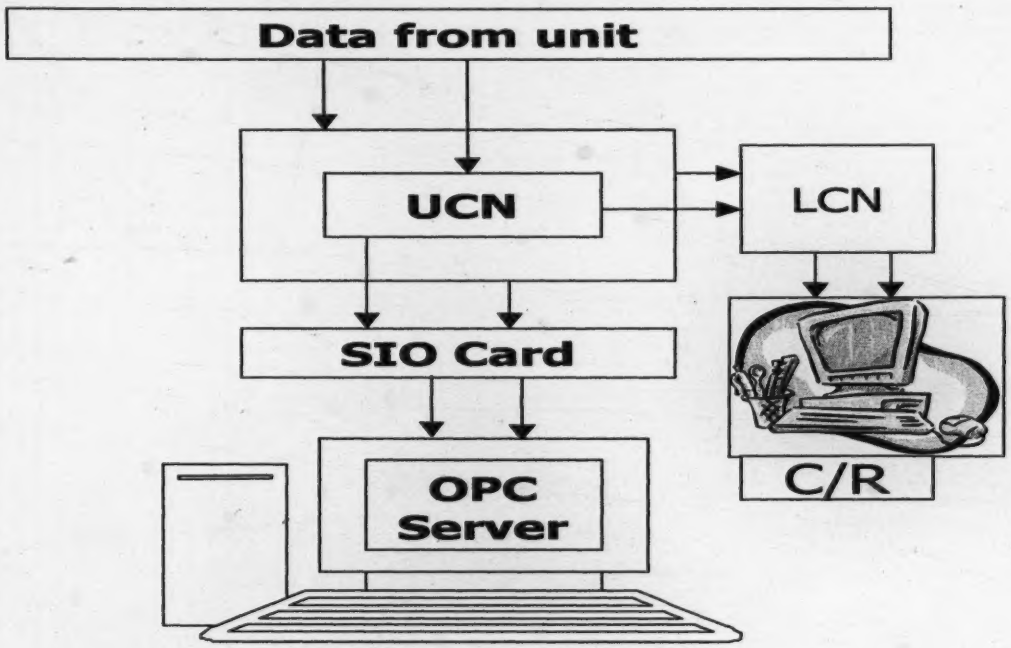


Figure 3.3: Schematic showing the placement of the Data Server on the existing DCS

3. The recycle component and the supercritical conditions make the process calculations for mass flowrate tedious and requires computing power.

Overpressure Relief Equipment: The high pressure section of the plant has isolation and dump control valves in case of emergency shut down. In addition, the piping and equipment have bursting discs which are designed to release the contents in case of over pressure. The reactor vessel is protected from over pressure, especially in case of decomposition by four bursting discs. These discharge into vent stacks which have automatic water injection systems to prevent auto ignition of the decomposition products.

3.3 Data Server and Variable Selection

In this application an OPC server(Object linking and embedding for Process Control) was used to access data from the Universal Control Network (UCN) of a Honeywell TDC3000 DCS. The OPC servers used Microsoft's OLE technology to communicate

with clients and permits standards for real-time data exchange between Windows based software applications and process control hardware. The data communication between the server and the DCS was via two serial RS-232 serial interface cards installed on one of the marshalling cabinets on the Universal control Network of the DCS (See Fig.3.3).

The polymerization reaction variables and the online product physical properties like Melt Index, Vinyl Acetate Wt.%, etc. are accessed using this system. Other variables accessed are the reactor process temperatures in various zones, the catalyst injection rates (these control the reaction temperatures), the reactor pressure, the feed temperature, the feed flow rate (derived from the secondary compressor variables), the product cooler variables, the recycle system variables, the propagation additives and the stirrer motor power. A total of 48 variables (also called tags) are stored every six seconds. In reality eight points are accessed in one second followed by the next set of eight points the next second until all the forty eight tags have been accessed and stored in a cycle of six seconds. The cycle is then repeated continuously. All the forty-eight tags could not be accessed at the same instant due to data transmission limitations from the DCS to the OPC server.

3.4 Special Events in Plant Operation

One special operating condition and two faults are summarized in this section. These special events will be discussed in subsequent chapters.

3.4.1 Cooler Cooking

The recycle stream from the high-pressure separator to the recycle gas coolers has entrained low molecular weight polymer (called low melt), which, upon cooling, forms a scale on the tube internal surfaces. With time, scale also forms on the suction and discharge piping of the secondary compressor, the inter-coolers, the feed gas coolers and the product coolers (See Fig.3.2). This leads to increased pressure drop and a higher fouling factor affecting both conversion and plant throughput. Cooler cooking is an exchanger surface cleaning operation which is done to revive heat exchanger efficiency by operating the exchangers at temperatures significantly higher than the

normal operating temperatures. At this condition the polymer and wax build up melts and is swept away with the feed or product gases, leaving the heat exchange surfaces clean. The product made during cooler cooks is off-grade material.

3.4.2 Reactants Decomposition

The lower alkenes are endothermic compounds; i.e. energy must be added to the component elements to produce the compound-this energy is +12.5kcal/mole for ethylene. Whereas propylene, another lower alkene, requires severe conditions to decompose explosively, ethylene decomposes under routine operating conditions (Britton and Wobser 1986). While ethylene undergoes a controlled exothermic reaction to form polymers, it may also undergo other exothermic reactions which may lead to a 'thermal runaway' and finally a decomposition reaction. A thermal runaway is defined as the increase of thermal rate acceleration of an exothermic reaction system from any temperature at which the rate of heat gain exceeds the rate of heat loss from the system. Alternatively, ethylene decomposes by two other methods, sudden compression, especially in the presence of impurities like diatomic gases, and by direct ignition.The decomposition initiating mechanisms are as follows:

- **Thermal Runaway:** Fire exposure, excessive temperature, stagnation, low-flow conditions, catalysis, excessive inlet temperature, heat of adsorption.
- **Direct Ignition:** Hot spot, flame propagation from associated equipment or piping, friction.
- **Compression ignition:** Trapped diatomic gas/presence of air, short compression time.

Numerous experimental studies have shown that thermal runaway decomposition is dependant on temperature, pressure and the geometry of the equipment in which the decomposition occurs. The locus of the critical temperature at which decomposition might occur is inversely proportional to the pressure of the system as outlined in the Frank-Kamenetski thermal analysis (Kanury 1977).

3.4.3 Loss of Reaction

The temperature of each zone of the reactor is controlled individually by regulating the rate of catalyst injection to that particular zone. The rate of catalyst injection in turn affects the polymerization rate proportionally which raises the temperature of the zone due to the exothermic nature of the reaction. Thermocouples installed in the zone will measure the temperature and will transmit the reading to the controller for action. In this reactor, we have a single loop feed back PID control strategy. Since the conditions in different zones may feed back on each other, good control of reaction temperature becomes imperative. It is well known that such exothermic polymerization reactors are open loop unstable reactors (van Heerden 1953, van Heerden 1958). These processes are characterized by:

1. High degree of conversion achieved after ignition.
2. Some feedback of heat along the reaction path.
3. Existence of three stationary states which satisfy the boundary conditions.

More recent research shows that more than three stationary or steady states are possible (Jaishinghani et al., 1977) and (Schmidt et al., 1981). Closed loop control strategies maintain the reaction temperature at the desired temperature, which is around the closed loop steady operating point. Loss of reaction in a zone is a condition where the reaction cannot be sustained in one or more zones as the controller is unable to reject a large disturbance and the reactor moves to a new stable but undesirable operating point. Two of the stationary points are stable. One of these has a low reaction temperature and a low conversion, the second has a high reaction temperature and near complete conversion. The intermediate state is an unstable one which is the operating stationary point. The cause of loss of reaction may be categorized as:

1. An unforeseen breakdown of the initiator or catalyst pump-this is the most serious disturbance and generally causes shut down.
2. A feed or initiator flow disturbance.

3. A disturbance in pressure control of the reactor which feeds back in the temperature loop.

In this thesis, PCA modelling is used for fault detection and can be used to predict the fault if it is slow to develop.

3.4.4 Grade Transitions

While it is preferred to have long campaigns for a particular grade, high inventory and less grade changes, the final operational policy depends greatly upon market forces. Therefore, operating multi-product plants is a demanding task and the goal is to achieve optimum capacity utilization; this is important to the profitability of the company. Conflicting interests of marketing, manufacturing, and plant control have to be balanced.

The complexity of operating procedures during grade change depends upon the difference in their recipes. Product changeover consists of partially overlapping activities, i.e. phasing out the previous product; adapting and changing over to the new product; and developing analytical procedures to determine the new product quality. Generally the grade change is achieved dynamically by continuously the catalyst, modifiers and other conditions. In certain cases, however, the recipes are so widely different, especially with regard to the catalyst type, that the product change requires stopping the reaction and then restarting. Process complexity is increased when another operation like a cooler cook is done simultaneously with the grade change to minimize downgrade material.

3.5 Concluding Remarks

In summary the process is very difficult to operate due to its dynamics and supercritical operating conditions. Incidents like breakdown of a catalyst pump leaves no time for the operators to take the spare pump in line and forces a short unit outage due to loss of reaction. There are multiple safety interlocks because of the critical conditions of operation which demands high standards of instrumentation maintenance for plant on-stream reliability. Operating complexity is increased due to grade slides and special operating conditions.

Chapter 4

Industrial Data Analysis

4.1 Introduction

The objective in this chapter is to explain the procedures adopted and applied consistently to data on the various products and operating conditions in this plant. The metrics used for assessment of abnormal situations are also consistently applied to various situations. Consistent assessments will aid industry in understanding the important and often difficult points involved in applying statistical tools to make process decisions.

Adequate care has to be taken to use appropriate tools for improvement in quality and productivity of the process. Since the thesis concerns the production aspects of LDPE and EVA, statistical process control(SPC) techniques have been chosen to monitor and model in order to develop control strategies in the future (Kim and Larsen 1998).

Modelling of this process is a complex task (Piovoso and Pearson 1992)

- Multiple grades are manufactured by the same process unit based on market forces.
- Unplanned disturbances occur related to the unit itself.
- Unscheduled changeovers or special operating procedures are used to revive operating efficiency.
- Modifications are sometimes done to the standard operating procedure for a grade.

It should be noted that the model developed is only applicable to the conditions under which the data was gathered. This has been experienced while doing this work and has been cited by other researchers (Piovoso and Pearson 1992, Kresta and MacGregor 1991). The following sections give a description of the tools and the application of these tools to plant data to create PCA models. These models are subsequently used for fault detection and diagnosis.

4.2 Data Preprocessing

Information was collected on the operating procedures and grade campaign production plans of the plant and a procedure was developed for preprocessing the data by grade or event.

The data pretreatment or preprocessing procedures essentially consist of selection of variables for monitoring, autoscaling and removal of outliers. The objectives while developing a standard procedure for data preprocessing were:

- Should be applicable and fairly easy to use by a process engineer.
- Should be necessary for the development of monitoring models.
- To be optimized to minimize computing costs.
- Should be adaptable and consistent.

of The selection of variables using the loadings matrix (refer to equation 2.4) could be done in a number of ways including(Kasper and Ray 1992):

1. We could examine the loadings of the first few components and look for clusters which signify highly correlated variables. One variable out of these clusters may be selected with little loss of information.
2. Variables with consistently low values of loadings may be dropped as their correlation with other variables is weak. These do not explain the process in any significant way.
3. Or in another approach, the variable with the largest loading in the first component is selected and the PCA modelling is repeated. The process is repeated until there are just enough variables to explain the process (Kasper and Ray 1992).

The first two selection methods from above were adopted for our case and 48 variables were chosen. Since the unit was generally run at capacity with the exception of cooler cooks and transitions, the mass flow rate of the reactants did not vary greatly during normal operation. A mechanistic model to evaluate the flow rate was specifically developed to study the effect of flow rate variation during cooler cooks and grade transitions.

The data was queried from the **OPC server** at the plant site using ***SQL*** (structured query language) program files. Data for the 48 variables (Appendix A) was accessed by the server in a sequential cycle of six seconds and stored in database folders, as explained earlier. This data was then queried for every 12hour shift using SQL program files, since all the transitions and special operations performed are logged by the shift personnel. By following a shift by shift wise query routine, it was easier to follow and understand the operations. The queried data was in 48 individual files for the 48 variables being accessed. Their total size was close to $440MB$ so they are compressed and stored in a user folder. Monthly data was queried in the above fashion and recorded on a data CD and taken to the university for further processing. Each data files had three pieces of information (See Appendix B):

1. Date and time stamp for the recording.
2. The process value recorded to the third decimal place.
3. The status of the tag or sensor being accessed, whether it was on or off.

A Unix based ***Perl*** script was then used to cut and paste the process variable value from the 48 files to a matrix format. This data matrix was then transferred to a *Windows* environment as a ***.txt*** file for further processing. The data matrix was then loaded to $MATLAB$ software and the following operations were performed in sequence:

1. The first operation was to eliminate data vector columns of: variables which had nearly constant values like controller set points and some controller outputs, variables with consistently low loadings identified from earlier modelling experiences, variables not relevant to the current grade. The controller outputs from only catalyst pumps that were on-line were kept.

2. The feed flowrate calculation was carried out using a secondary compressor model developed in *MATLAB* and was appended to the data matrix as a data vector.

3. The data matrix was then averaged over every six observations and then filtered using *EWMA* filtering.

4. The data matrix was then scaled so that each variable had zero mean and unit variance, except for gas environment analyzer process variables.

A number of *normal* or in-control steady-state regular operating data sets were preprocessed in the above fashion and appended to model the operation of a particular polymer grade. Since equipment changes (eg.catalyst pumps) lead to changes in the model (i.e. one equipment may have a different operating mean and variance than another), we only used data when similar equipment was operating. The PID tuning constants of the controllers also remained the same during the various grade runs. There is a gain scheduling built in for safeguarding against larger than usual plant disturbances. Adequate precautions were taken to reject data during abnormal disturbances and only 'steady- state' data was considered.

Preprocessing of data is crucial for the following reasons:

- PCA is purely a data driven technique.
- PCA based fault detection assumes steady state data for the model.
- PCA is scale dependant and therefore, data sets have to be scaled adequately.

Although we can manually detect outliers by inspection of plots, it is undesirable to do this because of the high volume of data involved. Therfeore, an EWMA linear filter was designed to filter raw data. The filter parameter value used varied based on the grade and the level of variance in the operation. The models obtained from filtered data have better resolution to special events, and minimize false alarms since noise is removed from measurements. Several months of data have been collected in various grades, cooler cook operations and other special events. The *normal* operation data set and scaling are crucial to the successful application of the procedure (Kresta and

MacGregor 1991). The following features are required for the selection of the data set and scaling:

1. The 'NOC' data set variation should be large enough to avoid false alarms. However, if the variance is too high then the model becomes 'thick' and becomes insensitive. It has been shown in this work that high variance models are used for fault detection in high variance operations and vice versa.

2. The scaling is done to give equal weight to all the variables for modelling. Auto scaling, that is, mean centering followed by scaling all variables to unit variance, is the most popular. However, variables with very small variance are not scaled to unit variance as they may become too large following scaling and may falsely dominate the model.

3. No extraneous variables are included since they may increase the threshold of the SPE plot (Kresta and MacGregor 1991).

4.3 Data Filtering

Process measurement noise may be electrically generated or process generated. Electrically generated noise is due to issues like cable shielding, grounding etc. and process noise is due to turbulence, multiphase flows, variations in mixing etc. (Seborg and Mellichamp 1989). Signal noise leads to reduced resolution and may cause false alarms in process monitoring schemes. These effects are reduced via filtering or signal conditioning before doing statistical modelling. Averaging over every six samples was carried out followed by exponential smoothing.

An exponential low pass filter was used to smooth the data sets by removing the high frequency fluctuations in the data. This filter can be represented by a first order transfer function or a first order differential equation,

$$y(t) = x(t) - \tau_f \frac{dy(t)}{dt} \tag{4.1}$$

where $x(t)$ is the measured value at time instant t, y(t) is the filtered value at that sampling instant and τ_f is the time constant of the filter. τ_f is selected based on the frequency range of noise associated with the measured variables. However, the

selected filter time constant should be smaller than the dominant time constant of the process. Here it is the residence time of the reactor which varies from 35 to 47 seconds for most grades and operating events. In the case of using the strategy for closed loop control the condition for choosing the filter time constant is,

$$\tau_f < \tau_{dominant} \tag{4.2}$$

Since the noise frequency range is an important consideration, a higher value of τ_f gives more smoothing. Though higher smoothing results in good resolution, we compromise on the prediction time of faults since a dynamic lag is introduced (Seborg and Mellichamp 1989). In PCA modelling the filter time constant depends on the noise associated with the measurements. After development of the PCA model based on a certain 'NOC' data set we cross-validate with another 'NOC' data set in the same polymer grade to check that the SPE plot 99% threshold is not violated more than the stipulated one out of every hundred samples. In the case that the violation is higher than the stipulated figure we increase the smoothing by increasing the filter time constant or increase the number of PC's in the model.

The digital form of the exponential filter is written using backward difference form for the derivative in equation 4.1, we get,

$$\frac{dy}{dt} \cong \frac{y_n - y_{n-1}}{\Delta t} \tag{4.3}$$

where n and $n-1$ signify subsequent sampling steps. On substituting equation 4.3 in equation 4.1 we get,

$$\begin{aligned} x_n &= \tau_f \frac{y_n - y_{n-1}}{\Delta t} + y_n \\ y_n &= x_n - \tau_f \frac{y_n - y_{n-1}}{\Delta t} \\ y_n &= \frac{\Delta t}{\tau_f + \Delta t} x_n - \frac{\tau_f}{\tau_f + \Delta t} y_{n-1} \\ y_n &= \gamma x_n + (1-\gamma) y_{n-1} \end{aligned} \tag{4.4}$$

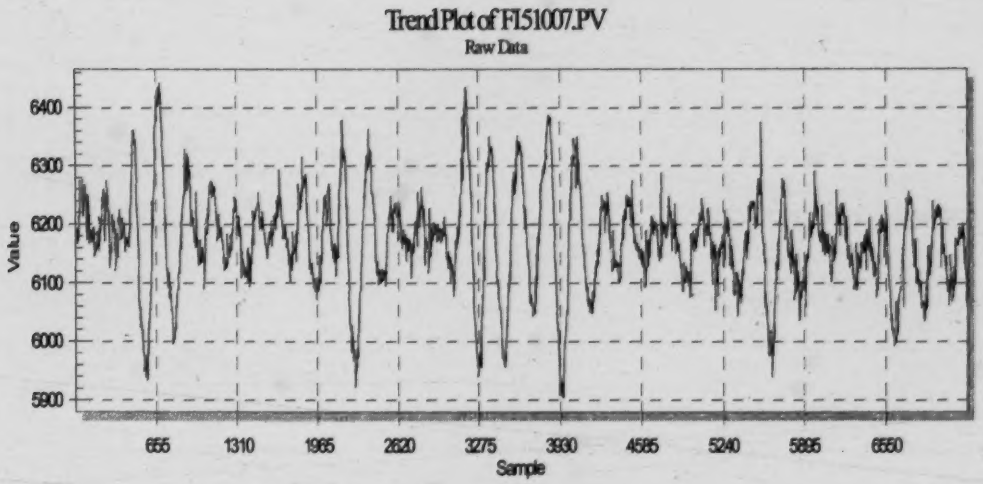


Figure 4.1: Raw Data Plot

where γ is defined as,

$$\begin{aligned} \gamma &= \frac{\Delta t}{\tau_f + \Delta t} \\ &= \frac{1}{\frac{\tau_f}{\Delta t} + 1} \end{aligned} \tag{4.5}$$

In our case $\gamma = 0.13$ to $\gamma = 0.15$ were used for the exponential filter, which give filter time constants in the range $\tau_f = 40$ to $\tau_f = 34$ seconds. For simplicity, this filter was applied to all data vectors, including the noisy flow variables, and may not be sufficient for products with higher variance. One reason for high variance is that the PID tuning constants of the controllers are the same for all grades. The exponential filter of the form given in Equation 4.4 has been used throughout this work for smoothing data before performing PCA modelling. Figures 4.1 and 4.2 show a comparison between raw and filtered data.

4.4 Secondary Compressor Model and Dynamic Flowrate Calculation

The high pressure process has a recycle stream of unreacted monomer which is cooled and recycled back to the secondary compressor suction. The secondary compressor

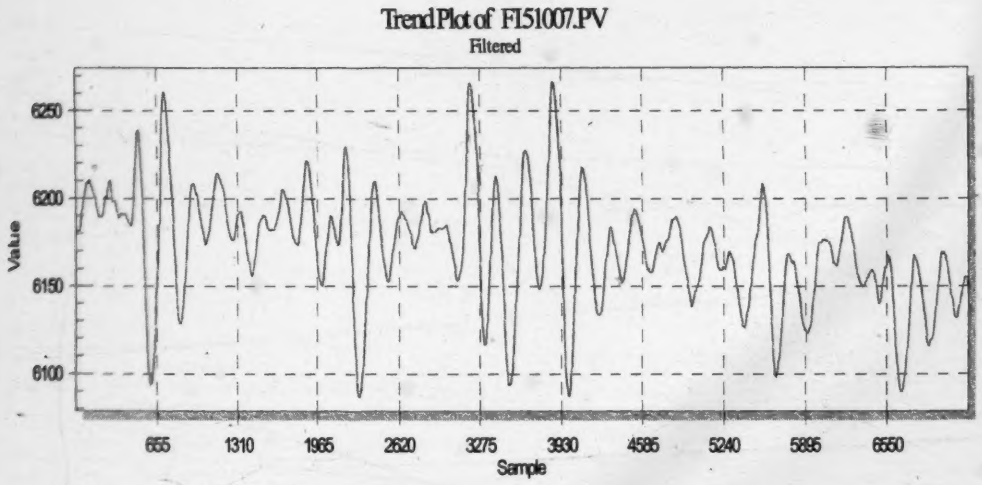


Figure 4.2: Filtered Data Plot

operates at fixed speed, and therefore, the suction and discharge conditions will determine the feed gas rate through the reactor.The feed gas rate is an important variable which affects product quality and plays a role in special events like decompositions and loss of reaction. A model of the compressor has been developed using thermodynamic correlations obtained from archives at AT Plastics Inc. Fig 4.3 gives the steps followed to determine the flowrate. A program was written in *Matlab* software to calculate the flow rate at every sample instant. The model is readily adaptable to an online model since it takes the data vector and the model and associated function files take less than 0.25 seconds for execution on an *Intel Pentium* $4, 1.4$ *Ghz* speed desktop machine.

4.5 PCA Modelling

A number of pre-processed data matrices for the same grade were combined or used to develop a PCA model. The data from different campaigns were merged to have a large representative data set. The modelling was done using a multivariate statistics and control performance assessment software called ***ProcesDoc (Matrikon Inc.)*** and cross checked using the *PCAGUI* in *MatLab PLS* tool box. The data was first plotted and combined with the plant log (history) of the operations to visually determine the

'steady state' data for the modelling. As explained in Chapter 3, the correlation matrix was estimated and an eigenvalues decomposition carried out to obtain the loadings matrix P_q. The success of the monitoring scheme depends on the following major aspects (Wise and Kowalski 1990):

1. The accuracy of the loadings matrix, P_q which must be estimated from data having good excitation or a high signal-to-noise ratio compared to the operation under consideration. The accuracy is improved with large data sets and preprocessing as explained in Chapter 4.

2. Only events similar to those used in the model can be compared with the 'normal' model. In other words, a global model may be too 'thick' to detect faults and a model developed from data with low variance or a low signal-to-noise ratio will give false alarms for an operation with high variance or a high signal-to-noise ratio.

The number of principal components (the number of eigenvectors in the loadings matrix) were elected by performing a scree test and in some cases confirmed by cross validation of the model with a 'normal' data set. Fig.4.4 is the scree plot for one polymer grade and shows the cut off line for the number of PC's adopted in the model.

Generally the eigenvectors associated with eigenvalues $\lambda_i \geq 1$ corresponding to seven PC's were chosen and have been found to be sufficient for our purpose. The cumulative variance with seven principal components in most grades was between 88 to 90% or these seven PC's described between 88 to 90% of the variance in the process. See Fig.4.5 and Table 4.1 for the same grade showing the cumulative variance and the loadings matrix, the associated eigenvalues and the tag configuration for the same grade respectively.

Additionally, the contribution of each variable to the individual PC's are depicted in the loadings plots(Figures 4.6 and 4.7) for the first two PC's of this grade. In this case, PC1 encompasses the feed conditions such as mass flowrate, secondary compressor variables (which essentially define the mass flowrate), feed gas pressure, the catalyst consumption and the feed temperature. PC2 captures all the reactor temperatures.

The covariance or correlation matrix, the loadings matrix, the associated eigenvalues and the tag configuration comprise the PCA model. The standard data set used for modelling and the standard statistics (the minimum and maximum values, mean and variance for each process variable) are stored. Data from a fault or a shift in operating point is compared to the standard set. The usefulness of comparing fault data to the standard data is explained in detail in the later chapters on fault detection and diagnoses.

Another useful tool used to analyze the data and relationships between variables is the *color plot of the variable correlation matrix* . It is a variant of the correlation matrix color plot and includes the correlation of between PC scores and the variables. It is developed by appending the scores columns to the data matrix and then evaluating the correlation matrix. This matrix has an additional advantage in that it shows the set of variables explained by the specific PC's. The correlation matrix is then plotted in color as shown in Fig.4.8. The first 40 columns are process variable correlations and the the columns 41 through 47 represent the correlations of the PC scores to the variables. The smallest and largest elements of the combined correlation matrix are assigned the first and last colors given in the color bar. Colors for the remainder of the elements in the combined correlation matrix are intermediate and are shown in the color bar. As expected there is no correlation amongst the PC scores. A detailed explanation on the application of this plot is given in the chapter on quality design.

For instance, the first principal component is positively correlated with the secondary compressor temperatures (suction and discharge variables 23 and 24), but is negatively correlated with the catalyst injection rates. This matches our operational knowledge of the process; that is, the higher the compressor suction temperature the lower the mass flow . Hence, we require lower catalyst injection rates to sustain the reaction. Catalyst injection rates, however, have minor correlation with the secondary compressor conditions. The correlations can generally be explained physically or via reaction kinetics.

4.6 Conclusions

Data preprocessing for development of statistical models is by far the most time consuming step. Therefore, it is of utmost importance that the steps involved for data pretreatment be chosen with care. *A priori* process knowledge for variable selection, statistical requirement of autoscaling were done alongwith data filtering. The flowrate variable in the data matrix helped in the resolution of process variables responsible for some process faults as presented in subsequent chapters.

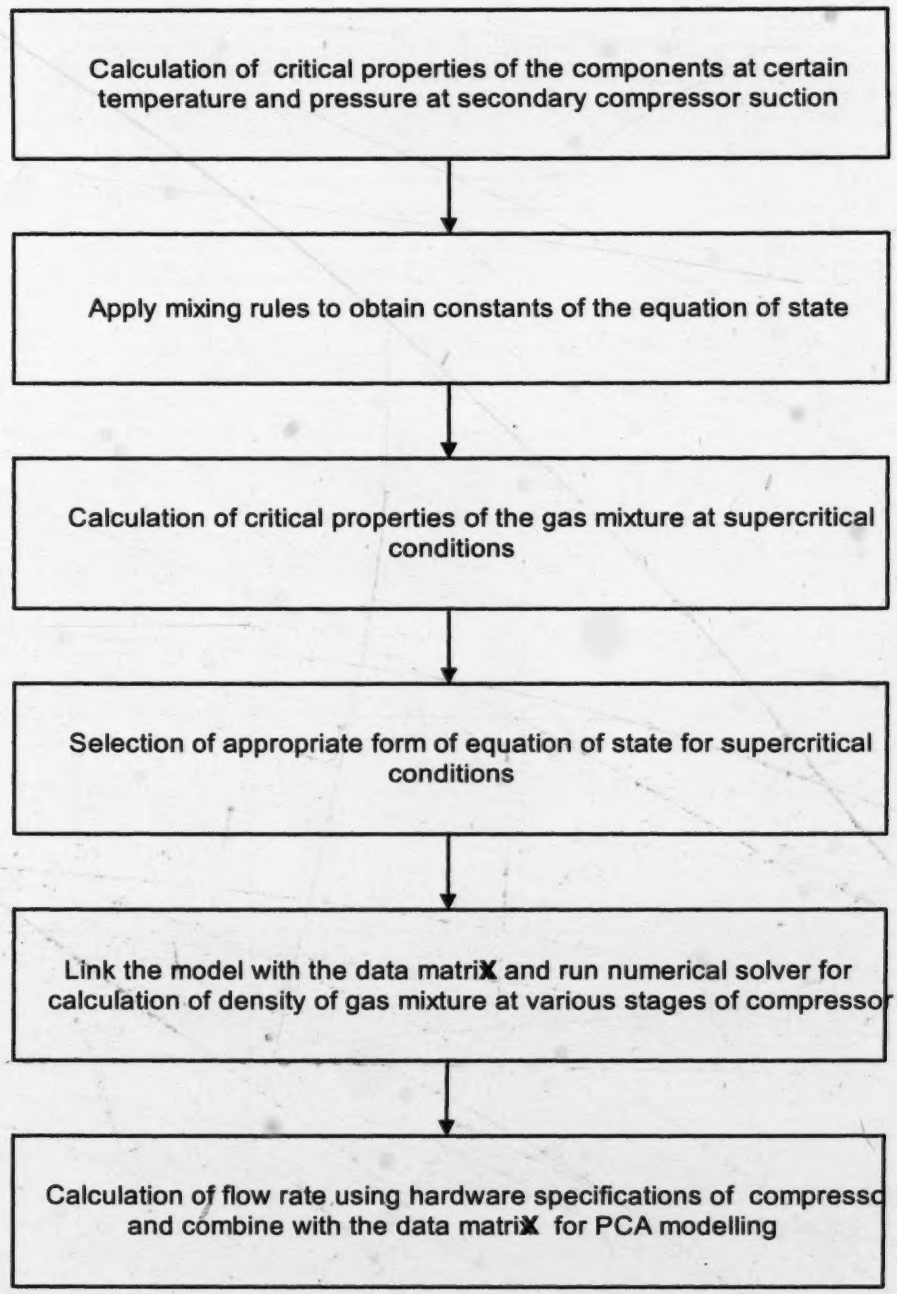

Calculation of critical properties of the components at certain temperature and pressure at secondary compressor suction
Apply mixing rules to obtain constants of the equation of state
Calculation of critical properties of the gas mixture at supercritical conditions
Selection of appropriate form of equation of state for supercritical conditions
Link the model with the data matrix and run numerical solver for calculation of density of gas mixture at various stages of compressor
Calculation of flow rate using hardware specifications of compressor and combine with the data matrix for PCA modelling

Figure 4.3: Secondary Compressor Model for Flowrate Calculation

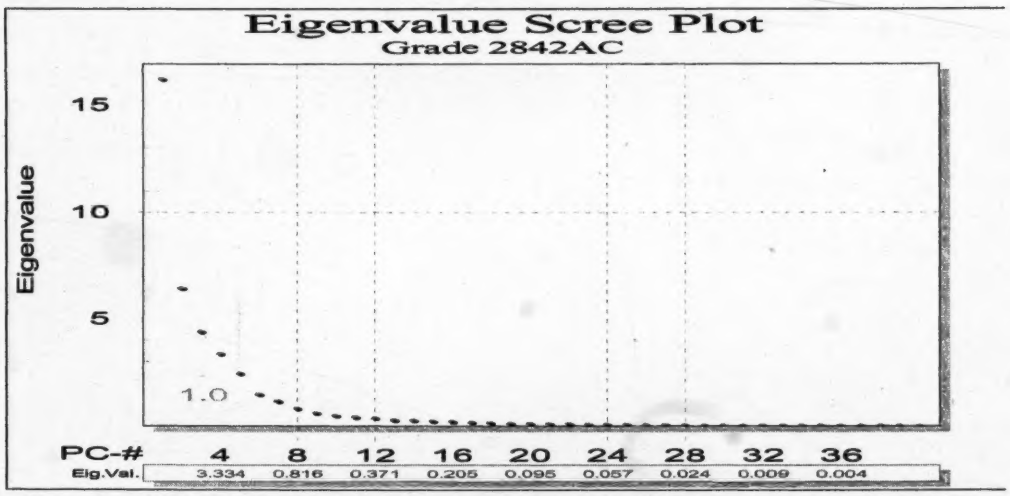


Figure 4.4: Eigenvalue Scree Plot

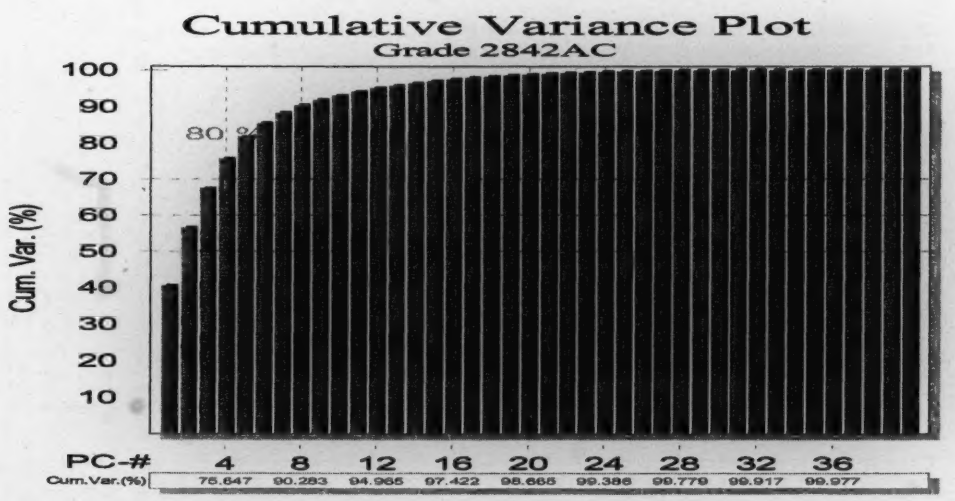


Figure 4.5: Cumulative Variance Plot: Seven PC's

S.No.	Variable	Eigenvalues	Cum. Variance	PC-1	PC-2	PC-3	PC-4	PC-5	PC-6	PC-7
1	Flowrate	16.16	40.39	0.79	-0.17	0.51	-0.06	0.06	-0.04	0.16
2	AI51001B.PV	6.40	56.39	-0.84	0.03	0.27	-0.30	-0.11	0.00	0.11
3	AI51001C.PV	4.37	67.31	0.77	0.05	-0.42	-0.22	-0.14	0.18	0.14
4	AI51001D.PV	3.33	75.65	-0.90	0.14	-0.22	0.00	-0.01	0.20	-0.02
5	AI53501.PV	2.43	81.73	-0.30	-0.03	0.25	0.55	0.25	0.60	0.19
6	AI53502.PV	1.46	85.37	0.44	-0.11	0.17	0.55	0.38	0.47	-0.03
7	FI51007.PV	1.15	88.24	-0.83	0.06	0.28	-0.05	0.14	-0.06	-0.04
8	FI54010.PV	0.82	90.28	-0.87	0.00	0.11	0.32	0.19	-0.03	-0.08
9	FI54121.PV	0.60	91.78	-0.82	0.03	0.23	-0.13	0.09	0.14	0.20
10	II52005.PV	0.49	93.00	-0.37	-0.05	0.43	0.15	-0.01	-0.20	-0.57
11	LIC52001.PV	0.41	94.04	0.65	0.15	-0.59	-0.21	-0.10	0.10	0.15
12	PI51135.PV	0.37	94.96	0.89	-0.12	0.07	-0.20	-0.08	0.07	-0.02
13	PI51176.PV	0.28	95.67	0.88	-0.16	0.37	-0.20	-0.06	-0.01	0.03
14	PI52051.PV	0.27	96.33	-0.34	0.19	-0.69	-0.40	-0.23	0.26	0.01
15	PI52054.PV	0.23	96.91	0.66	0.05	-0.17	0.02	0.24	0.00	-0.25
16	PIC52021.OUT	0.20	97.42	0.65	-0.15	0.39	-0.44	-0.14	0.26	0.02
17	PIC52021.PV	0.17	97.84	-0.03	0.16	-0.09	-0.24	0.23	0.31	-0.61
18	PIC52022.PV	0.13	98.17	0.84	-0.19	0.42	-0.22	-0.10	-0.04	-0.03
19	SI53016.PV	0.10	98.43	-0.94	0.02	0.12	-0.07	0.04	-0.02	0.05
20	TI51015.PV	0.09	98.66	-0.18	0.23	-0.85	-0.08	0.13	-0.05	0.01
21	TI51021A.PV	0.09	98.88	0.92	-0.14	0.24	0.11	0.07	0.06	0.04
22	TI51154.PV	0.08	99.08	-0.30	0.18	-0.69	0.46	0.24	0.18	0.01
23	TI51158.PV	0.07	99.24	0.73	0.08	-0.50	0.07	0.03	-0.09	0.05
24	TI51164.PV	0.06	99.39	0.83	-0.10	-0.04	-0.20	-0.14	0.22	-0.25
25	TI52022.PV	0.06	99.52	0.06	0.78	0.26	0.31	-0.33	0.02	0.06
26	TI52024.PV	0.05	99.64	0.47	0.69	0.13	0.13	-0.43	-0.06	-0.04
27	TI52026A.PV	0.03	99.72	0.07	0.80	0.07	0.41	-0.34	0.01	-0.03
28	TI52029A.PV	0.02	99.78	0.12	0.84	0.11	-0.25	0.41	-0.01	-0.01
29	TI52031A.PV	0.02	99.82	0.09	0.64	0.05	-0.48	0.47	-0.06	0.17
30	TIC52005.PV	0.02	99.86	-0.81	0.12	-0.26	-0.32	-0.19	-0.05	-0.21
31	TIC52021D.OUT	0.01	99.89	-0.84	-0.04	0.34	-0.24	-0.11	0.19	0.13
32	TIC52021D.PV	0.01	99.92	0.11	0.81	0.11	0.23	-0.42	0.05	0.00
33	TIC52025D.OUT	0.01	99.94	-0.86	-0.05	0.32	-0.20	-0.11	0.25	0.10
34	TIC52025D.PV	0.01	99.95	0.16	0.81	0.11	0.23	-0.44	0.05	0.00
35	TIC52027D.OUT	0.01	99.97	-0.10	-0.23	0.06	-0.50	-0.32	0.48	-0.13
36	TIC52027D.PV	0.00	99.98	0.10	0.85	0.19	-0.10	0.19	0.16	-0.11
37	TIC52028D.OUT	0.00	99.99	-0.84	0.06	-0.09	-0.40	-0.29	-0.01	0.02
38	TIC52028D.PV	0.00	99.99	0.19	0.82	0.15	-0.25	0.39	0.02	-0.04
39	TIC52030D.PV	0.00	100.00	0.22	0.63	0.20	-0.47	0.45	-0.10	0.13
40	TIC52034.PV	0.00	100.00	0.98	-0.07	-0.10	-0.01	0.00	0.10	0.03

Table 4.1: Loadings and Tag Configuration of the model with seven PC's

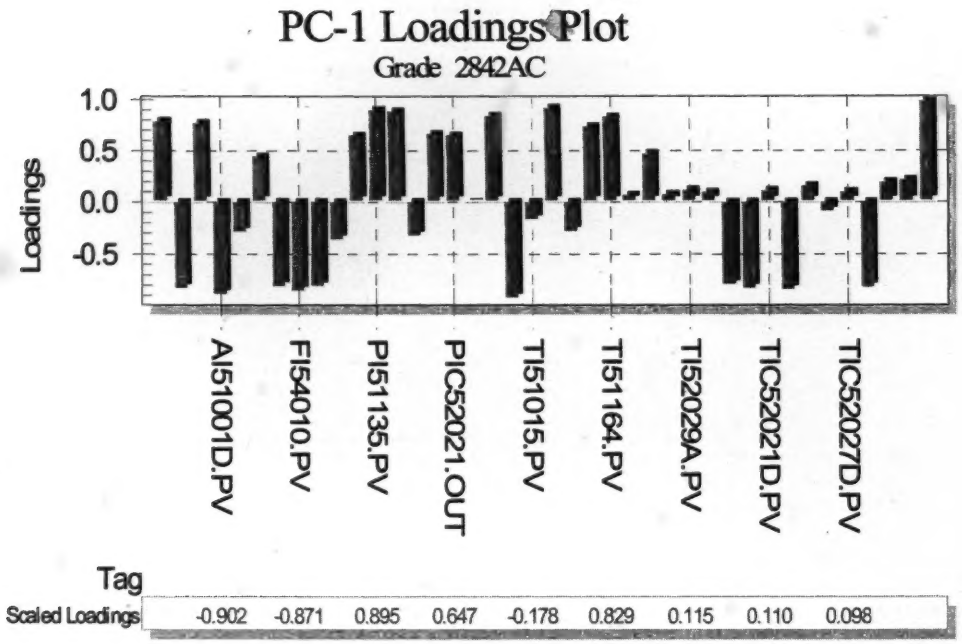


Figure 4.6: Loadings Plot:PC 1

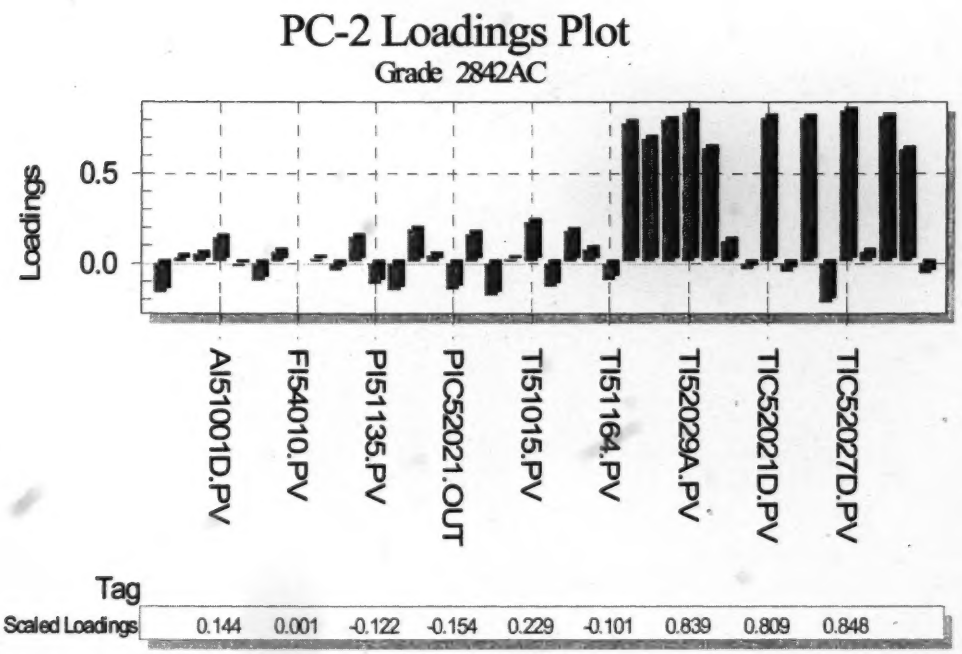


Figure 4.7: Loadings Plot:PC 2

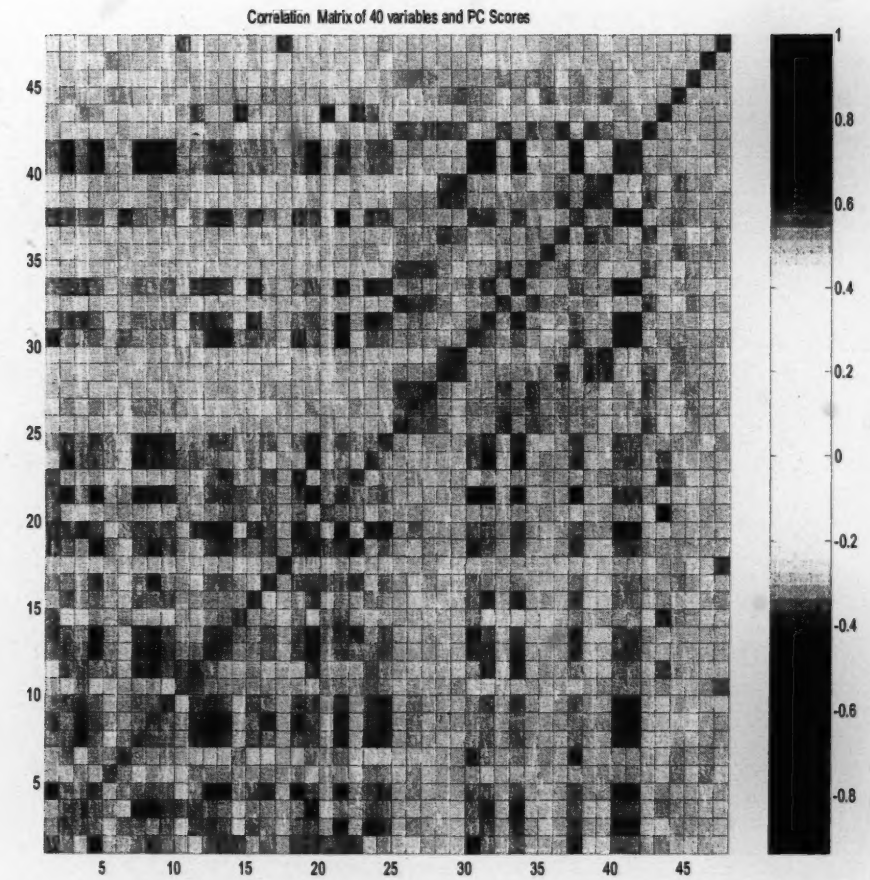


Figure 4.8: Correlation Matrix

Chapter 5

Fault Detection and Diagnosis: Decomposition

5.1 Introduction

This chapter explains the procedure employed to simplify and understand the process data and give conclusions from an off-line multivariate PCA analysis related to data from a decomposition (often termed as decomp in industry) during a campaign in the $18xx$ class of products, which are 18% Vinyl Acetate content copolymer.

Multivariate Statistical Process Monitoring, Fault Detection and Diagnosis using Principal Components Analysis (PCA) was demonstrated by Jackson, 1980. It was shown that these methods were suitable to analyze highly correlated process data. The advantages of using these methods are their capability to handle large data sets, typical in the chemical process industry, and dimension reduction via monitoring certain key statistics as in Shewart charts. While the theoretical development of Principal Component Analysis was done by Pearson, 1901 and became a part of the multivariate statistical tools, the application of the PCA based fault detection and diagnosis method began only in the last two decades. Wise, 1991 and MacGregor and his co-workers (Kresta and MacGregor 1991, Kosanovich and Nomikos 1994) have done pioneering work on PCA and PLS (Partial Least Squares) as applied to monitoring and control of the process industry. The techniques have also been demonstrated on the standard shell control problem for process monitoring and control by Kasper and Ray, 1992. Although we know the underlying principles in the unit processes and operations in the chemical process industry, it is daunting to set up a mechanistic model

of the same for monitoring and control; hence,the increasing popularity of the PCA and PLS based fault detection and diagnoses. However, an important pre-condition for diagnosis of root cause/s for a fault is sound process knowledge.

In this thesis we study the process of ethylene and vinyl acetate monomers undergoing exothermic reactions to form polyethylene and ethylene-vinyl acetate copolymers. Contaminants like oxygen and hydrogen react with the monomers in other exothermic reactions to create side products. Alternatively as explained earlier in Chapter 3, §section 3.4, ethylene and its mixtures may be directly decomposed at local hot spots or by sudden compression in the presence of impurities like diatomic gases. Such reactions may lead to a thermal runaway and finally an explosive decomposition (Britton and Wobser 1986). Thermal runaway is the main mechanism by which decomposition occurs in these high pressure LDPE reactors.

Decomposition of ethylene and its mixtures has been researched widely both at laboratory and industrial scale. This research has lead to safer and better design of equipment and procedures in the high pressure technology area. To avoid decompositions, the polyethylene industry has proprietary operating procedures or 'recipes' based on thermodynamic considerations and operating experience. To reduce the incidence of decompositions, strategies such as tighter automatic controls, mechanistic models for multivariable controllers and model predictive control have been implemented. All of these are geared towards tighter control at the 'recipe' operating point. As mentioned previously this knowledge, even if it has marginal success, is closely guarded for commercial competitiveness. Therefore, it is important to develop simple data based strategies which are easily implementable with a reasonable degree of accuracy and success to avoid decomposition and other faults.

During the campaign of the 18% VA class of products in October 2001, a decomposition was experienced at AT Plastics $R-V$ unit. This decomposition event took place while a routine cooler cook (refer to Chapter3, §section 3.4.1) operation was taking place. The decomposition resulted in a rapid rise in temperatures and reactor pressure and the reactor's safety bursting discs gave way. This rapid increase to abnormally high temperatures and pressure is typical once decomposition reactions have been initiated since the decomposition reactions are strongly exothermic. Although routine, cooler cooks are considered as special operating conditions since the reactor

is not at the stable operating point during the heating cycle but returns to it during the cooling cycle. To avoid temperature runaway, the changes in operating conditions should be small and the controllers should be well tuned.

As a result of the decomposition, unit $R-V$ was shut down with the automatic safety interlocks. The operating data for the campaign was later queried from the server to perform a detailed multivariate statistical analysis of the event.

5.2 Model Development

To detect a fault, the process operations data is compared with a 'normal' operation model. Since the decomposition took place during a cooler cook operation, statistically a different zone of operation than the general production conditions, a typical 'normal' operation model would lead to misleading conclusions. Following the philosophy highlighted by other researchers (Kresta and MacGregor 1991), an appropriate model using data only from similar events was developed to study the cause of decomposition.

A basic statistical comparison tells us that the 'normal' operation model has lower variances than those cooler cooks. Therefore, the 'normal' model ellipsoid is 'thinner' and will show the whole cooler cook operation as a violation of the 99% threshold. See Fig.5.1 for a score plot in the same grade which shows the 'normal' operation as a dense central scatter and the cooler cook as a deviation from the 'normal' operation regime. Figure 5.2 shows univariate trend plots of the compressor temperatures indicating that they are highly correlated with each other. The peaks in the figure correspond to the planned increases in temperature during cooler cook. It is these temperature excursions along with other changes in the plant which show up as violations of the 'normal' operation regime in Fig.5.1. Thus our 'normal' model must be representative of the cooler cook operation.

To develop a model for the decomposition analysis, *data from six cooler cook operations* in the same class of polymer grades with essentially *the same feed mix* were chosen and combined to form the representative data set. Process data from six cooler cooks which were done during the three days prior to decomposition were used to develop the 'normal' model. The cooler cook operation data which ended in

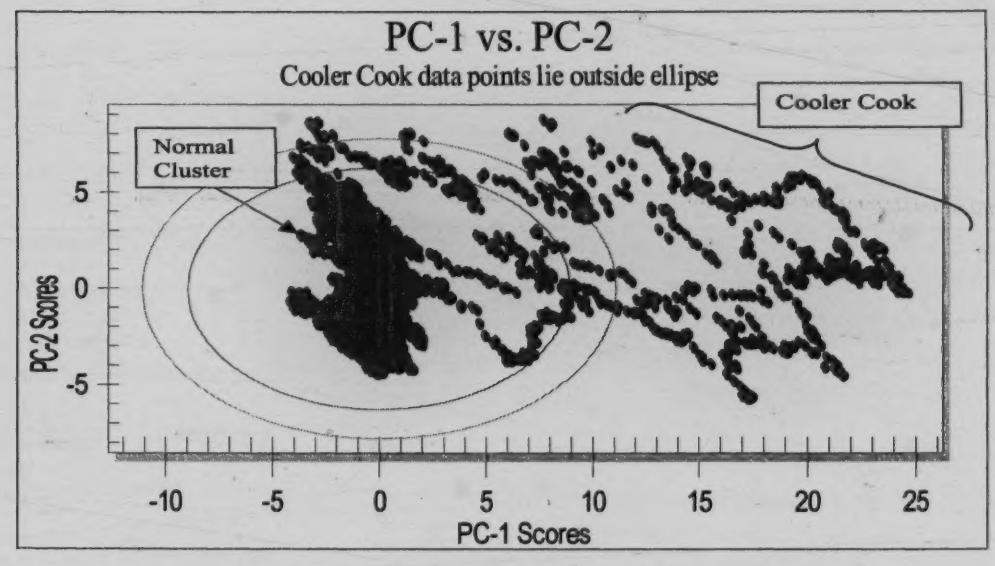


Figure 5.1: Cooler Cook comparison with normal operation

a decomposition was cross-validated with the 'normal' model.

Every cooler cook operation lasts for 1 to 1.5 hours and has been explained in detail §Section 3.4.1. The data from six cooler cooks was preprocessed using the steps outlined in the §Sections 4.2 and 4.3. Out of the 61 variables filtered and processed, 35 process variables were considered significant and used in the model. These included the analyzer readings of the gas composition, the electrical power to the stirrer motor, the secondary compressor variables defining the feed flowrate, the reactor temperature and pressure variables and the conditions of temperature and pressure in the high pressure separator and recycle gas coolers. The physical product quality variables of Melt Index and % Vinyl Acetate content were dropped since the online analyzers are recalibrated during cooler cook operations.

The advantage of an off-line analysis is that we can use process knowledge to remove known abnormalities from the representative data set. Thus variation due to random disturbance is dropped from the data set and only variation due to common cause is used in the model. This is done by plotting every variable to obtain a better understanding of the process variation. A part of the data set may have to be dropped if a special operation is done which is *not a normal procedure* during a cooler cook.

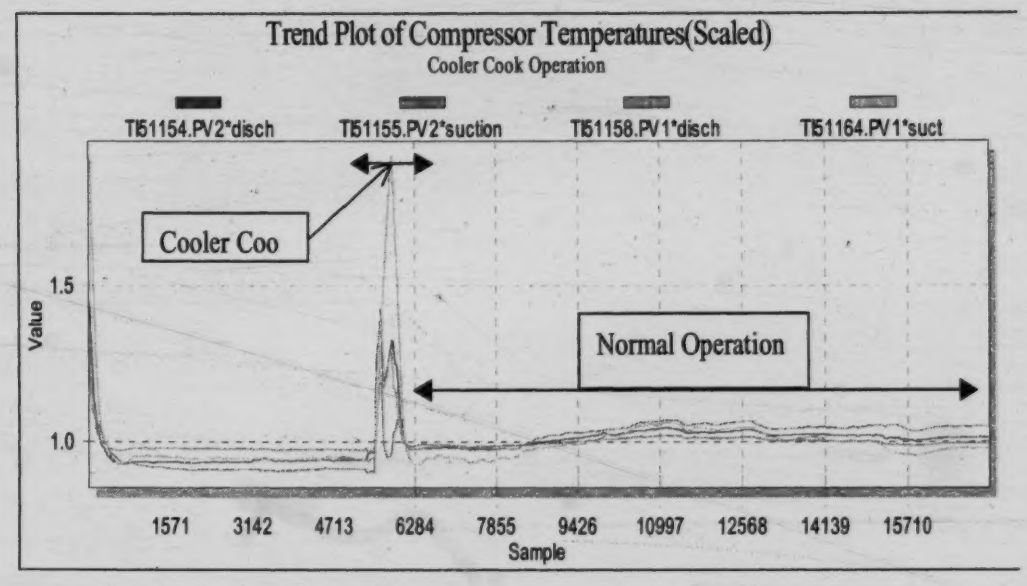


Figure 5.2: Cooler Cook comparison with normal operation trends

There may be abnormalities due to a malfunction in a control loop or equipment. We may also drop the variable altogether if through process knowledge ,it is believed that the variable has only remote possibility of being a cause of the fault. This is done to minimize false alarms and increase resolution for root cause diagnoses of the fault.

For example see figures 5.3 and 5.4 show two process variables with 'abnormal' trajectories in the sixth cooler cook. If these are in included as part of the model , the SPE value will be high at these times and the threshold limit will also increase thus reducing the resolution of variables responsible for the fault. However, in all such cases, process knowledge is the principal factor which helps us decide the list of variables which may be eliminated. In this case after choosing the 35 critical process measurements, data considered 'normal' was used to develop the model.

Fig. 5.5 shows the SPE of the same data used for development of the model by appending six cooler cooks during the same campaign. Since the cooler cook is a special operation there is a lot of variance in the SPE value.Fig. 5.6 shows the $2D$ scores plot which also shows the high variance of the cooler cook operation since the points are in a dense cluster as is usually seen for a 'normal' production campaign. See fig. 2.2 for a normal operation $2D$ scores plot.

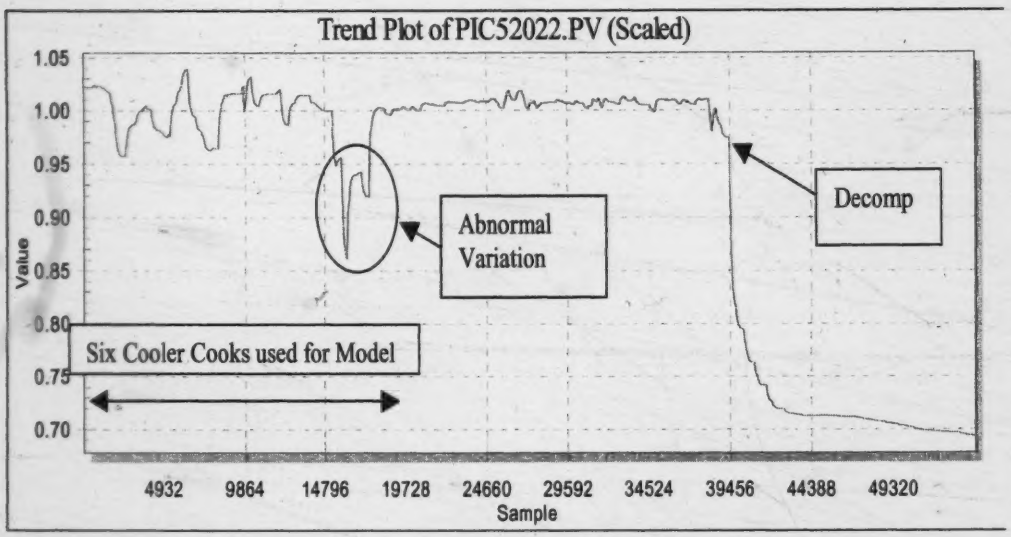


Figure 5.3: Plot showing abnormal variation of High Pressure Separator pressure affecting model threshold

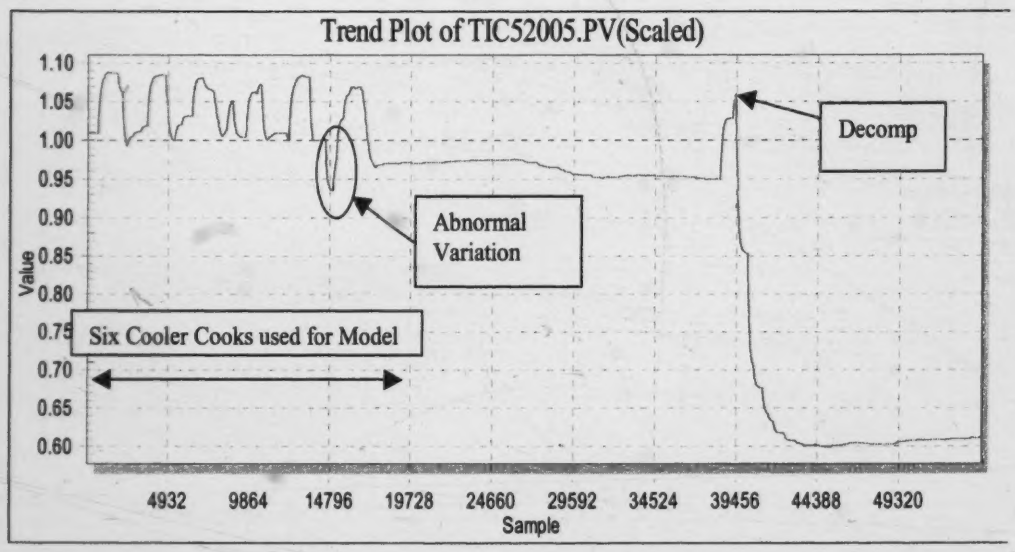


Figure 5.4: Plot showing abnormal variation of Product Cooler outlet temperature affecting model threshold

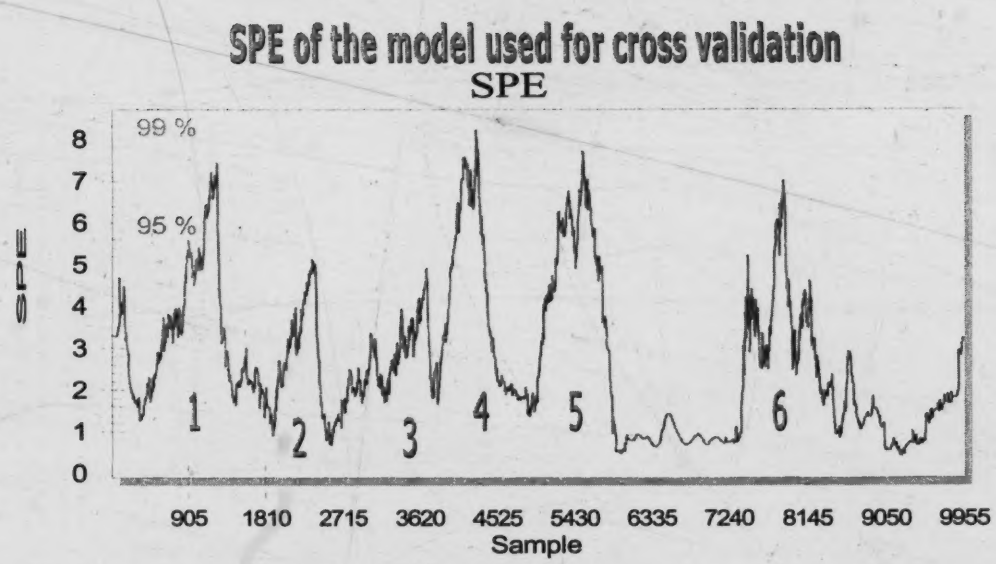


Figure 5.5: SPE of Decomp Model showing appended data from six cooler cooks

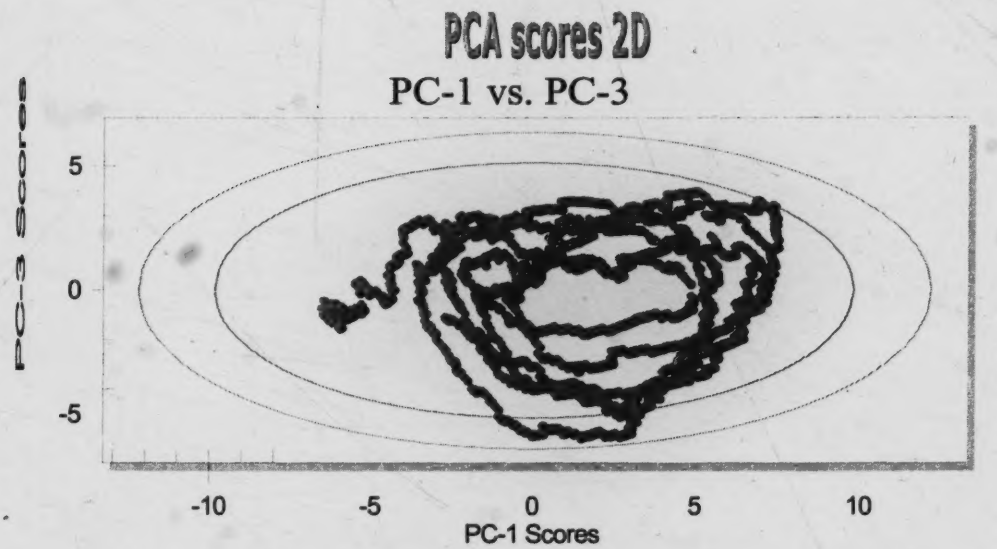


Figure 5.6: Decomp Model Scores

PC	Associated Eigenvalue	\% Cumulative Variance Explained	Variables Explained
PC1	16.92	48.34	Feed flowrate, Fresh VA flowrate, CO2 content,Reactor Stirrer current, Extruder speed, Secondary Compressor temperature and pressure conditions, Product cooler conditions,Catalyst Injection rates, Feed temperature, Reactor pressure control valve opening
PC2	9.53	75.56	Ethane content, Reactor feed gas pressure, Reactor pressure,HP separator pressure, Reactor lower Zones temperatures
PC3	4.22	87.6	Zone 1 temperatures
PC4	1.43	91.7	Gas environment analyzers

Table 5.1: Decomp Model PC's and their significance

The model captured 92% of the process variance with only 4 PC's. The eigenvalue scree plot and the cumulative variance are shown in the figures 5.7 and 5.8. A description of the 4 PC's, associated eigenvalues and the variables explained by each PC are given in the Table 5.1, and the loadings table is given in Appendix C.

5.3 Detection and Root Cause Diagnosis of Event

The break down of the correlation structure amongst process variables is termed as 'failure' of the process model. The SPE may violate the 99% confidence limit due to special events like decreasing the pressure of process equipment below the 'normal' operating limit. A fault, however, is generally detected when many variables contribute to violation of the confidence limits of the SPE plot.

Cross validation of the whole set of operation data, for the day on which the decomposition occurred, was done with the developed model using the chosen 35 variables.

In Fig.5.9 the SPE value exceeds the confidence or threshold limit, during the final cooler cook. The SPE value continued to violate the threshold limits This

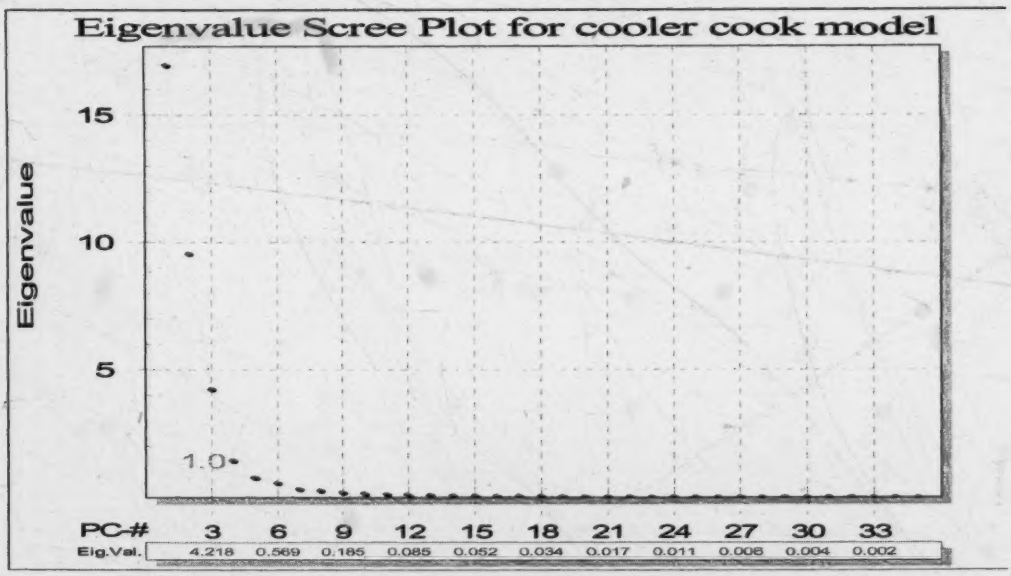


Figure 5.7: Decomp Model Eigenvalue Scree Plot

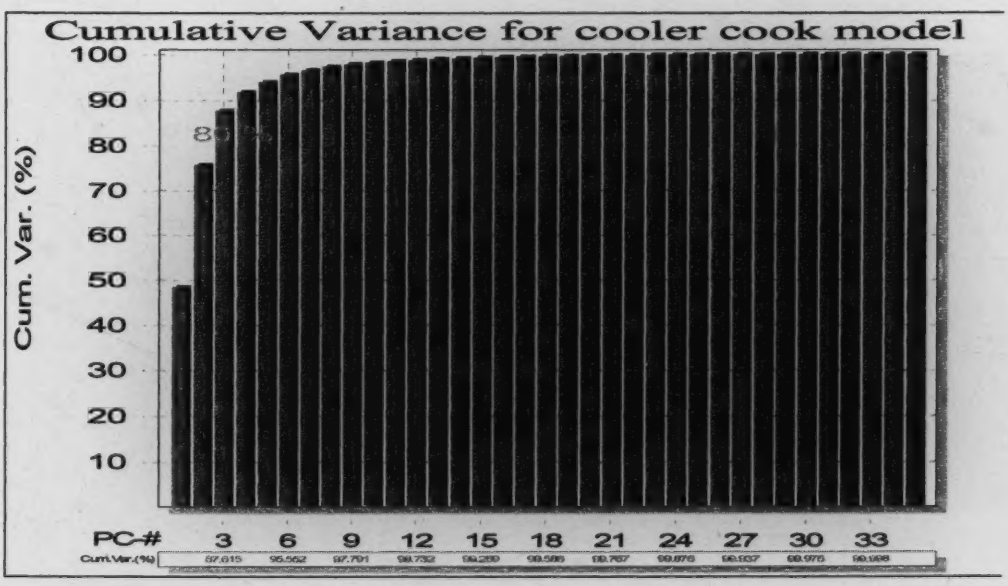


Figure 5.8: Decomp Model Cumulative Variance captured by PC's

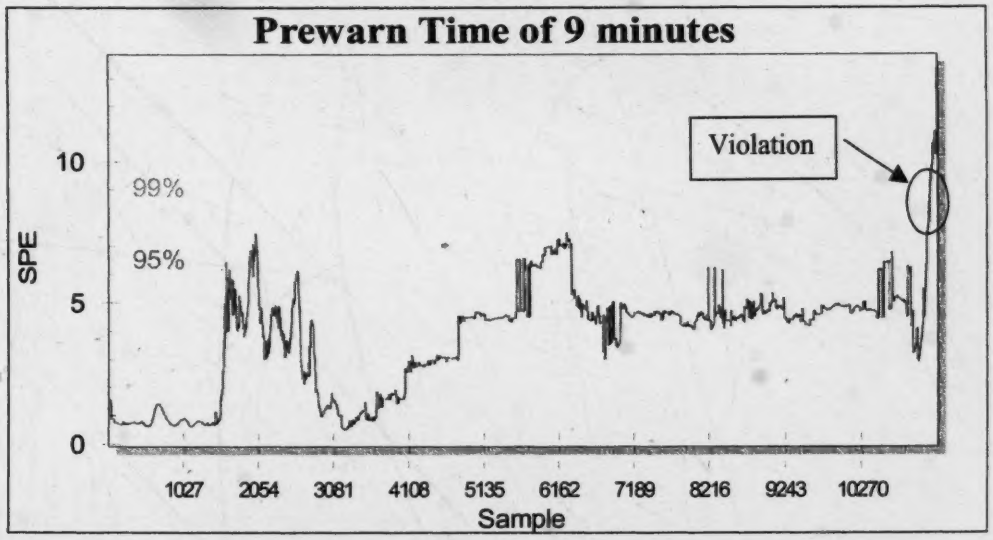


Figure 5.9: SPE Plot showing violation of threshold during seventh cooler cook

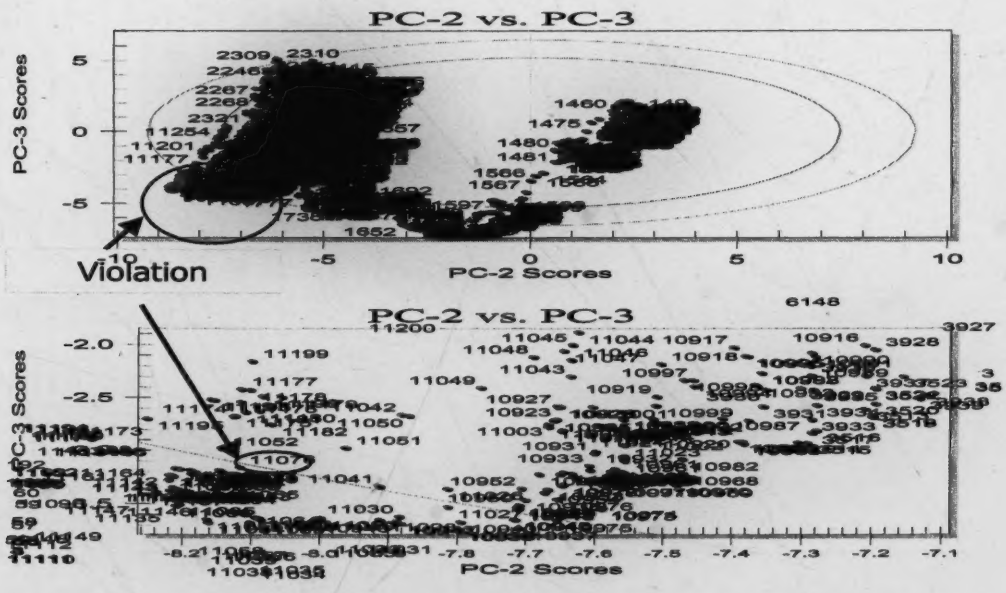


Figure 5.10: (a)Decomposition: Scores plot showing the violation of the threshold ellipse. (b) The lower figure is a zoom-in of the area where the violation occurs

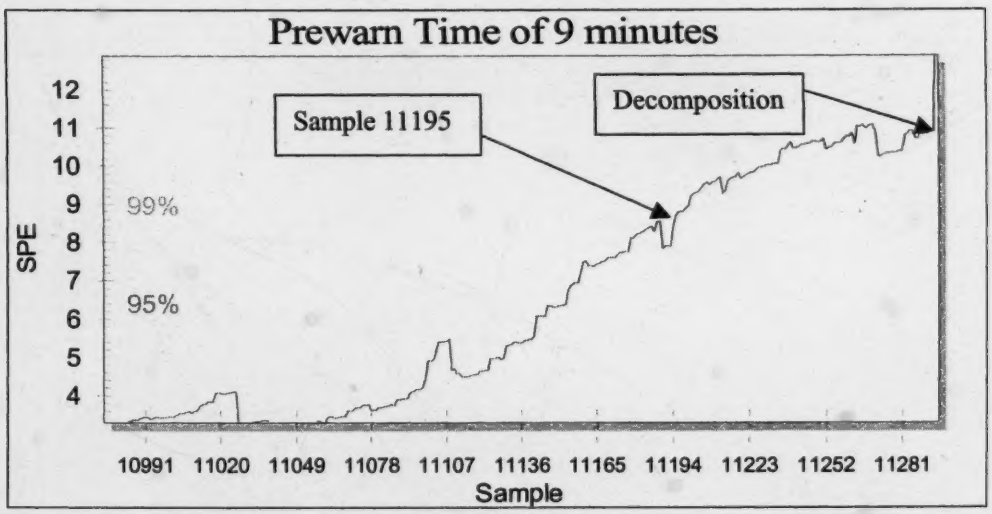


Figure 5.11: Zoom-in of the area where violation of the SPE threshold occurs and the point of decomposition

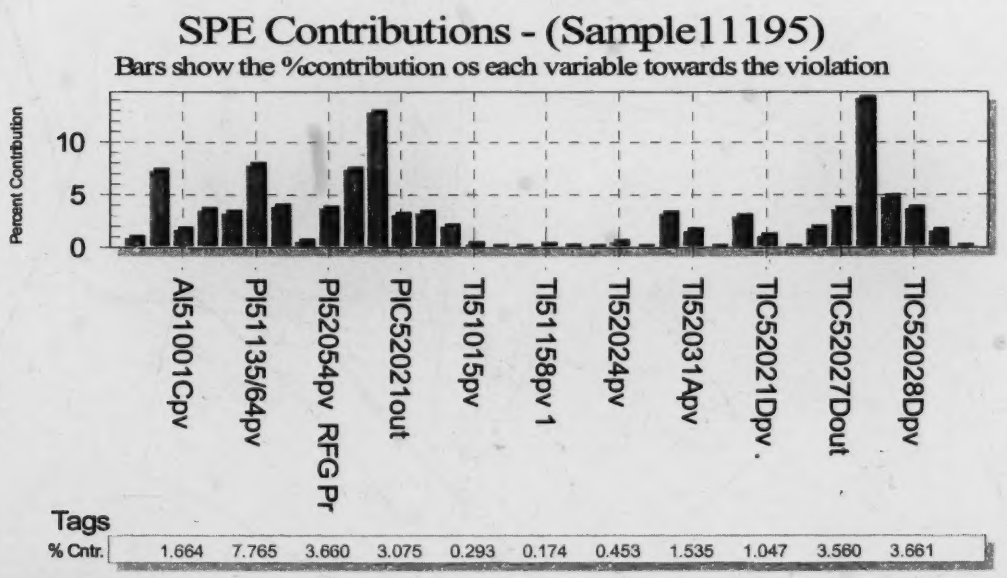


Figure 5.12: SPE Contribution Plot:Bars show the %contribution of each variable towards the violation

cooler cook had the quality of a sustained violation of the threshold limit, until the decomposition,i.e. its 'fingerprint' was different from the earlier cooks in the same grade.

The next step was to plot the contribution of each variable to the SPE value. Using this technique we isolate the variables responsible for the SPE violation or fault. The percentage contributions of the all the process variables in the model to the SPE violation of the 99% threshold ,can be plotted for each sample instant of the violation as shown in the bar plot of Fig. reffig:decompcontribution11195. Successive plots like Fig. reffig:decompcontribution11195 will point out the variables which are the largest contributors to the SPE violation.

Another way to find the variables responsible for the violation would be to note the contribution of each variable at every sample instant the SPE plot is above the 99% threshold. Thereafter plot the SPE contributions versus time as shown in the Fig.5.14.This technique has shown to promise for online implementation. Since the variables which have a sustained high percentage contribution towards the fault are most responsible, we may eliminate the other variables from the contribution plot unless the variable increases in importance as time proceeds. The contributions of all the 35 variables are plotted versus time in Fig.5.14and plot in the time series sense for all the 35 variables and Fig. 5.15 shows the top five contributors.

Decompositions has been widely researched in the high pressure polymerization industry but the kinetics are still not well understood. The following observations were made about the character of the decomposition analyzed in this work,

1. This decomposition occurred due to a combination of a number of variables being out of normal operating regime. This clearly why we need a multivariate analysis to detect faults.

2. The temperature of zone 2 of the reactor was consistently higher than normal operating levels and the pressure of the reactor was also high. The reactor pressure set point was also a high contributor to the SPE plot;it had been stepped up continuously during the cooler cook. Zone 2 temperature and reactor pressure were the highest contributors to the fault. The reactor internal check post-decomp also confirmed that the decomposition was initiated in zone2 of

the reactor.

3. The fourth highest contributor was the secondary compressor suction pressure which relates to the mass flowrate through the reactor. The flowrate model showed that while the flowrate was not as low as for the other cooler cooks, it did a contribute to the fault. See Fig.5.13 for a plot of flowrate versus sampling instants.

4. The fifth largest contributor was the ethane content of the system. Ethane is one of the products of a decomposition reaction (Britton and Wobser 1986) in addition to carbon, hydrogen and methane. Though it is the result of a decomposition, when the ethane level starts to increase, it may be an important indicator of an impending fault. However,the purge rate from the unit was not available and we could not confirm that the increase of ethane beyond normal levels was due to decomposition.

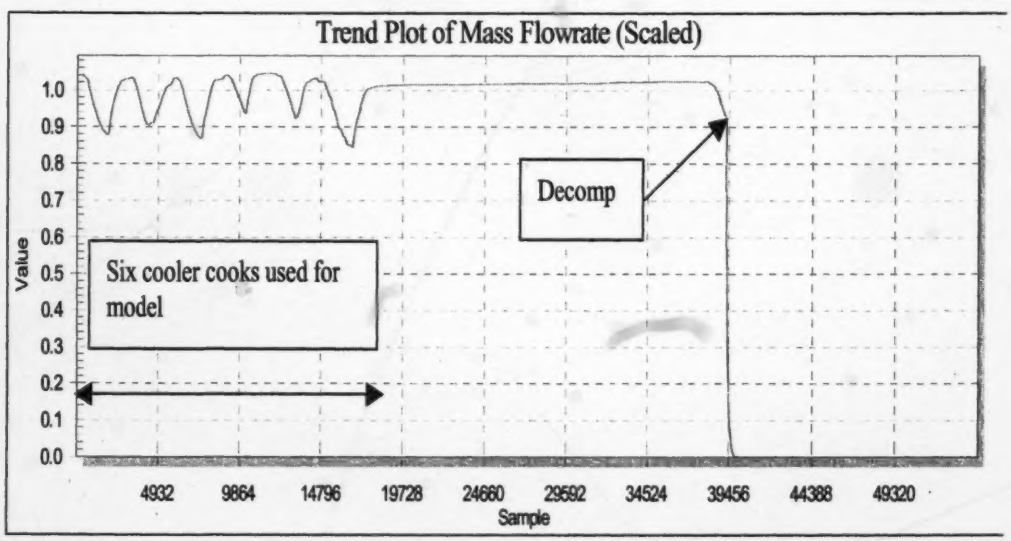


Figure 5.13: FLOWRATE in kg/hr versus sampling time for the six earlier cooler cooks and for the final cooler cook which ended in a decomposition

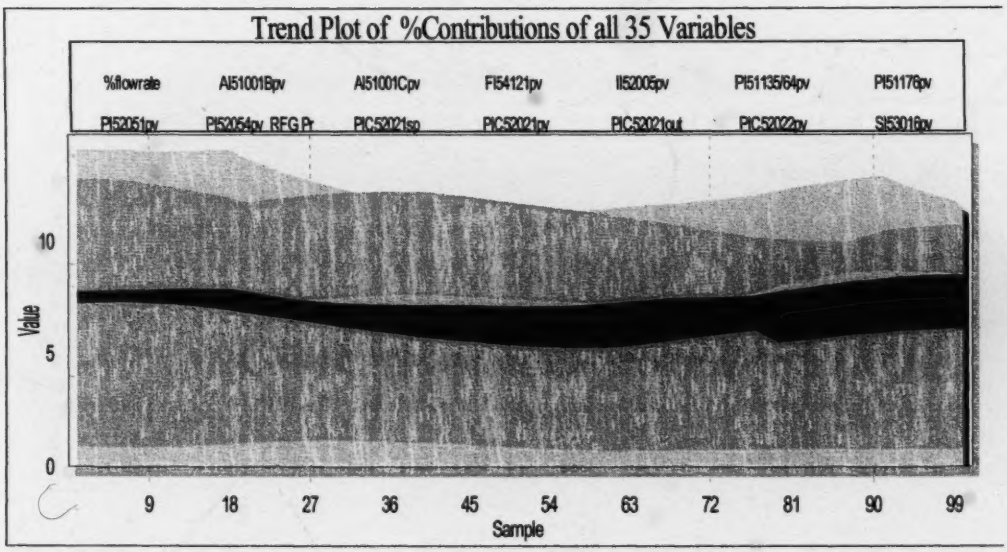


Figure 5.14: Trend Plot of percentage contributions of all 35 variables during the violation of the SPE threshold.It is hard to distinguish the top five contributors from such a combined plot.

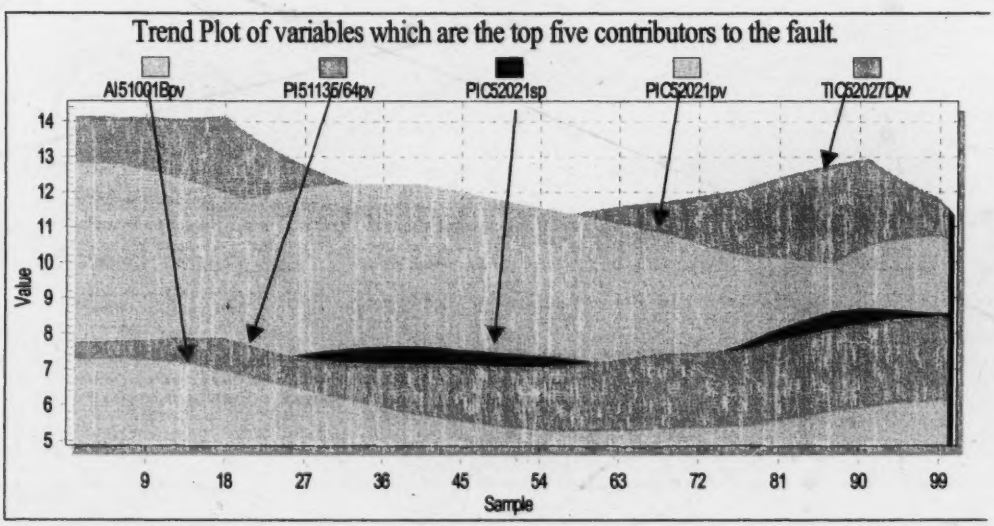


Figure 5.15: Trend Plot of the top five contributors to the violation of the SPE threshold.

5.4 Conclusions

The conclusions from the PCA analysis are:

1. Faults can be predicted using PCA based monitoring schemes although it is not necessary that a prediction will end in a decomposition. This technique can detect breakdown of normal variable correlation structure; i.e. variables do not have the same relationships as during normal operation.

2. Many faults are multivariate phenomena. As seen from the univariate plots, none of the of all the 35 variables showed any remarkable departure from their normal levels. Thus, univariate monitoring techniques may be used for event detection but are less useful for faults.

3. A reasonable prediction time is possible and corrective action can be facilitated by the variable contribution plots.

4. Even if PCA is not implemented online, it remains a potent analytical tool for offline process diagnosis.

5. Only similar situations can be compared. The PC scores plot of the decomposition model showed high variance;therefore,recursive techniques will have to applied for online implementation of the scheme.

6. The six earlier cooler cooks did not end in an undesirable situation while the seventh cooler cook did cause a decomposition and unit outage. Therefore, automation of cooler cooks will help standardize such operations and prevent faults.

Chapter 6

Fault detection and Diagnosis: Loss of Reaction

6.1 Introduction

Ethylene polymerization is exothermic in nature and the temperature increase across the reactor is a measure of the extent of conversion of ethylene monomer to polymer. The temperature increase across the reactor is limited by the spontaneous decomposition temperatures of ethylene and its comonomer. In continuous polymerization, a thermal equilibrium is maintained by removing heat at the same rate as is generated by the reaction. The heat removal is achieved by adding cold feed gas, which is generally at a steady mass flow rate. The polymerization is controlled by regulated addition of polymerization initiators commonly called catalyst. In effect the exit temperatures of the reactor become measures of the extent of polymerization and hence recipes will require that the temperatures be controlled within tight boundaries at a given reactor pressure and catalyst mix. It is important to understand that the temperatures in various zones of the reactor are highly correlated and the catalyst pumps injection rates in the various zones are also highly correlated. These results have been proved by doing correlation studies which are described in the next chapter. High correlation amongst the temperatures in the different zones of the reactor assists prediction of loss of reaction since an abnormal shift in the operating curve of one pump will directly influence the speed of the lower zone pump. This will lead to a break down in the normal correlation structure and through process monitoring, we will see an increase in the SPE and it will detect an event. This event may lead

to fault.

Loss of reaction in polymerization reactors is a condition in which the reaction temperature in the different zones cannot be sustained and this is an indication of low reactant conversion. This condition may start at the first or top zone and spills over to other lower zones since there is an equilibrium between all exit and inlet temperature conditions. In such conditions the reaction shifts to another stable operating point with very low conversion values. This situation is fairly common in the polymerization industry. Any knowledge extraction leading to further insight into the phenomenon of loss of reaction using general operation data is of immense value to the industry.

The reasons for loss of reaction have been explained in some detail in §Section 3.4.3 and are reproduced below:

1. An unforeseen breakdown of the initiator or catalyst pump. This is the most serious disturbance and generally causes a shut down.

2. A feed or initiator flow disturbance.

3. A disturbance in pressure control of the reactor which feeds back in the temperature loop.

The catalyst pumps are gravity fed from catalyst mixing tanks and storage vessels. The pumps themselves are high pressure, double-acting reciprocating units driven by individual hydraulic units giving a pulsating flow. The stroke length is controlled by the temperature controller of that zone; i.e. the catalyst injection rate is controlled by the dedicated closed loop temperature controller of the zone into which the catalyst is injected. Loss of reaction due to pump breakdown occurs because of many reasons:

1. An increase in the leaks through seals can affect the pump stroke length required to maintain temperature.

2. Catalyst in the seat of check valves can lead to valve malfunction which will also affect the stroke length.

3. Abrupt failure of a part, an interlock trip or power outage.

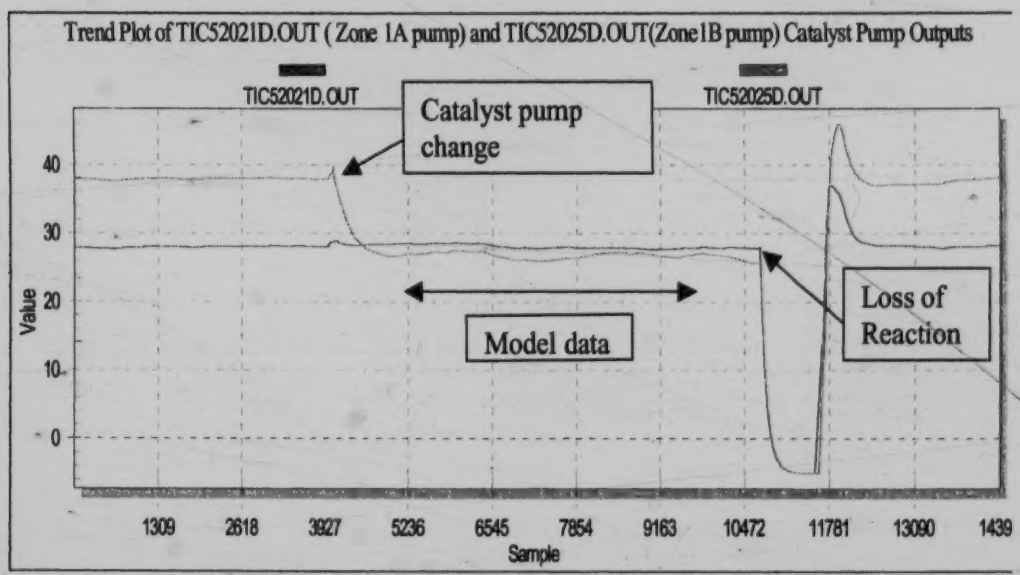


Figure 6.1: Trend Plot of Catalyst Pump Outputs in Zone1B showing Pump Changeover

Loss of reaction due to the first two causes occurs slowly and there is good potential it will be predicted by PCA techniques, enabling corrective action by operators. Abrupt failures may be detected but there is no time to avert a fault.

6.2 Model Development

A case study was performed on loss of reaction in zone $1B$ of the reactor in a 28% vinyl acetate class of product campaign. The fault occurred due to pump failure. Each zone of the reactor is has one operating catalyst pump and is backed up by one spare pump. It may be pertinent to point out that due to critical operating conditions and cyclic stresses experienced by the catalyst pumps, they have comparatively higher breakdown incidences. Depending on the running schedule of pumps and the preventive maintenance carried out on them, their operating curves can change over time. After five days of production with a particular combination of catalyst pumps, one pump was changed over previous to the loss of reaction breakdown. See Fig.6.1, the spare pump taken in line has a smaller stroke length to sustain the reaction temperature in the same zone. The stroke length of Zone1A pump is not affected after

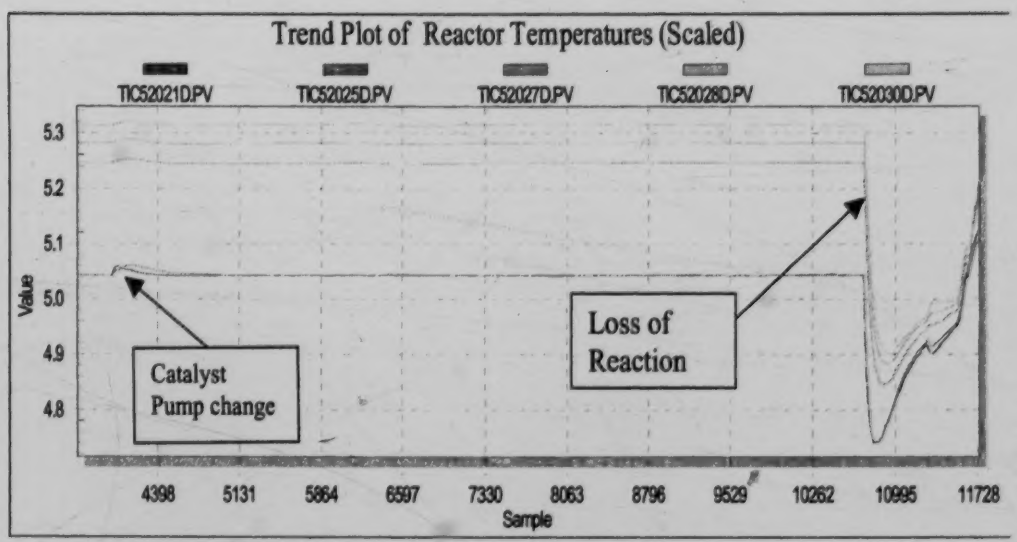


Figure 6.2: Trend Plot of Temperature Loops of Reactor

the catalyst pump in Zone1B is switched. Therefore, if the pumps are switched in any zone during a campaign there will be a mean shift in the operating point of that pump and will be shown as a bias in the SPE plot. To avoid this problem, data selected for the PCA model used the same set of pumps as when the loss of reaction occurred. Another way to deal with pump switching is to carry out mean correction for the data values for the pump output whenever pumps are switched. This correction may be applied with online implementation using pre-developed PCA monitoring models. Despite precautions taken by operations staff when a catalyst pump is changed, the reactor temperatures are still disturbed as seen in the Fig.6.1. Note that after the change-over, the temperatures are fairly steady until the loss of reaction. Additionally it should be highlighted that the output of the pump of zone1B, TIC52025D.OUT, does not show any remarkable change and a loss of reaction is experienced without any pre-warning based on a univariate analysis of the critical variables. However, a multivariate PCA analysis of the operation data gave a prediction time close to 40 minutes which could be used for corrective action.

For the model development, the data taken only while the same set of catalyst pumps were operating as those at the time of the loss of reaction was used. A total

of 4250 observations of 37 variables were used (See Fig.6.1). The chosen variables were the same as those in the decomposition model with additional online physical properties of the product(See Appendix D for the list of variables and their associated loadings in the model). The standard steps as outlined in Chapter 4 for data pre-processing were followed and a PCA model was developed. Using 5 Principal Components of the model explained 88% of the variance of the data. See Figures 6.3 and 6.4 for the scree plot and the cumulative variance captured by the PC's respectively. A description of the PC's and the variables which they describe are given in Table 6.1. This description is derived from the eigenvalues and the magnitude of the loadings associated with each PC.

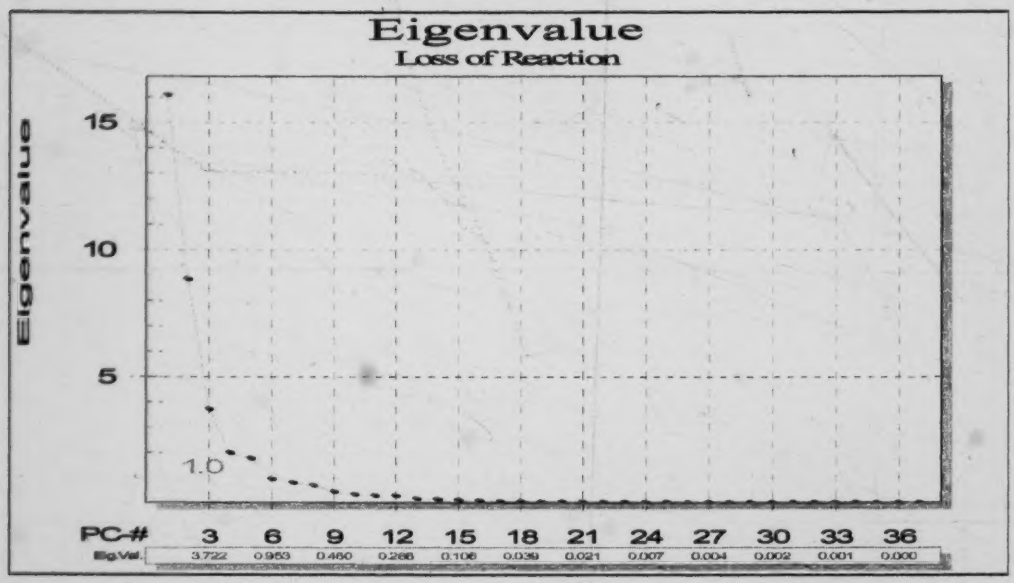


Figure 6.3: Scree Plot:Five PC's selected for Loss of Reaction model

6.3 Detection and Root Cause Diagnosis of Event

During the production of 28% VA class of products, an incidence of loss of reaction was chosen for detailed analysis. Reactor Zone1B experienced loss of reaction due to its catalyst pump failure leading to a cascade effect in the lower zones and a short unit outage. The pump failed due to something referred to as 'ball-over' in industry

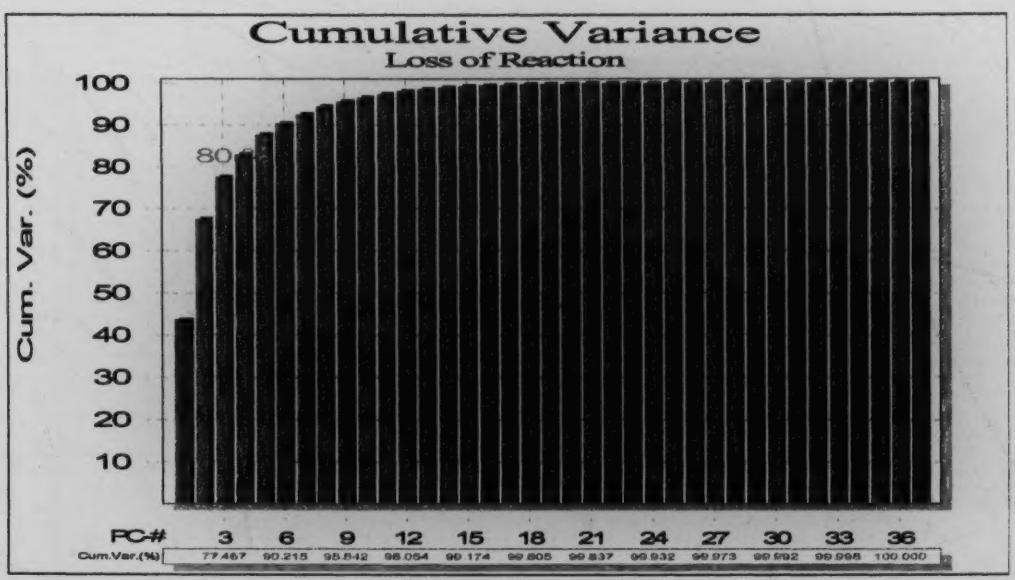


Figure 6.4: Plot showing cumulative variance captures by Five PC's selected for Loss of Reaction model

parlance which means that initiator deposited in check valves causing a malfunction. There was a gradual reduction in the catalyst flow rate until the flow ultimately stopped. Since the fault occurs over a long time, prediction of an impending fault should be possible.

Operation data for the day of the loss of reaction due to this pump failure was cross validated with the model developed in §Section6.2 . Hotelling's and SPE plots were obtained as shown in Figures 6.9, 6.10 and 6.10. Sample instant 1 to the final sample instant 10707 at which the failure occurred were considered for cross validation. It was observed that the threshold was violated at sample 9619 (See Fig. 6.10)and the corresponding variable contribution plot (see Fig.6.8 showed that the main contributor was the mass Flowrate evaluated from the secondary compressor variables as outlined in §Section 4.4. The univariate plot of the compressor interstage pressure shown in Fig.6.15 shows the abnormal cycle which caused other variables of the secondary compressors to exit their normal operation zones. This combination caused the violation of the SPE 99% confidence limit. One example is the higher temperature at the second stage discharge when the interstage pressure dipped causing

PC #	Associated Eigenvalues	Cumulative Variance	Variables Explained
PC1	16.1	43.5	Feed flowrate, Gas environment ethane content, Product MI and Vinyl Acetate content, Stirrer motor current, Secondary compressor variables, Reactor pressure, HP separator pressure, Product cooler outlet temperature, Reactor Zone1A catalyst pump output, Feed gas temperature.
PC2	8.8	67.4	Product cooler inlet temperature, Reactor Zone1B and lower zones catalyst pumps outputs, Reactor Zone1B and lower zones temperatures, Zone
PC3	3.7	77.5	Reactor Zone 1A and 1B temperatures
PC4	1.99	82.8	Gas environment CO_2 content, Return gas cooler outlet temperature
PC5	1.8	87.6	Extruder speed, Gas environment CO_2 content, Return gas cooler outlet temperature

Table 6.1: Table showing PC#,associated eigenvalue, variance captured,variables covered by the various PC's

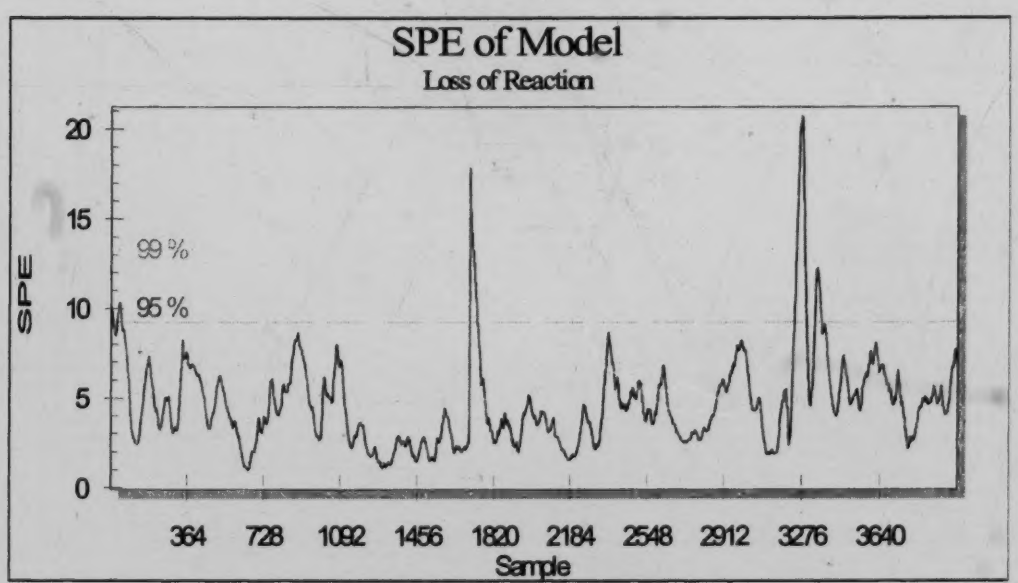


Figure 6.5: SPE Plot of Model

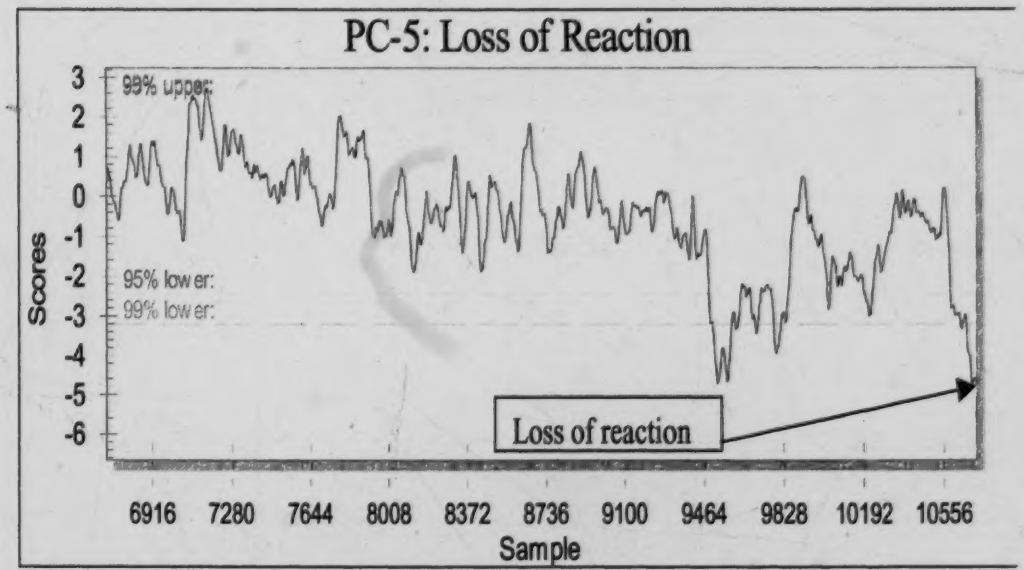


Figure 6.6: PC5 violation of threshold:Loss of Reaction

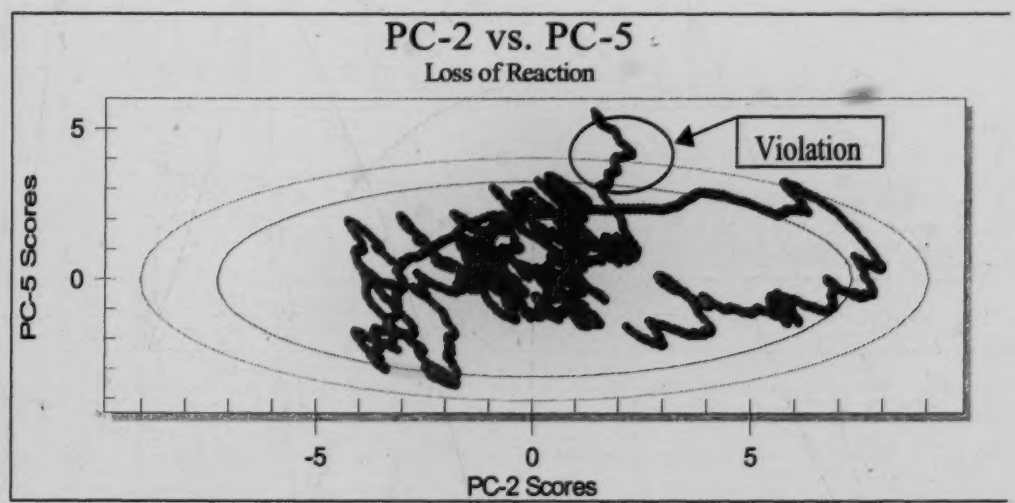


Figure 6.7: Score Plot showing violation of threshold:Loss of Reaction

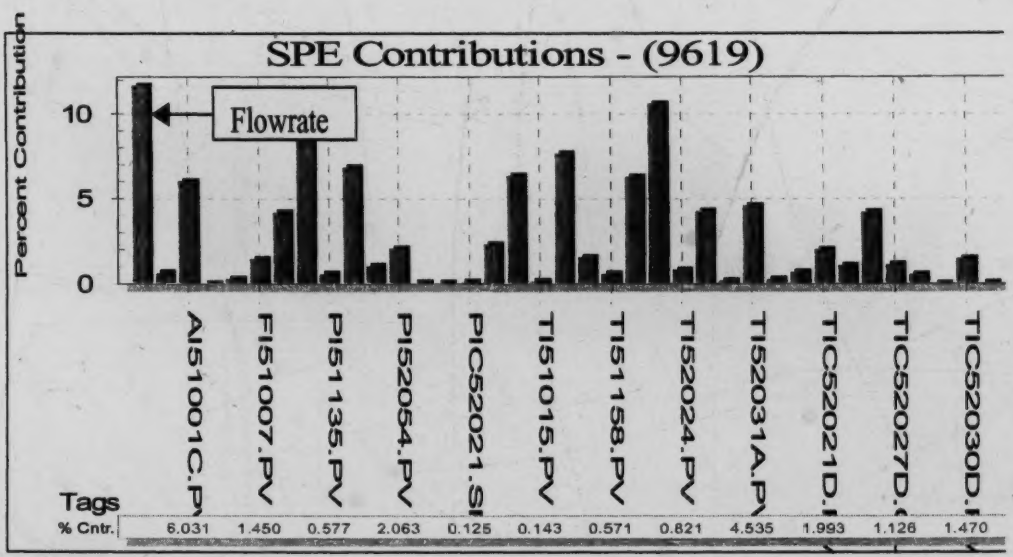


Figure 6.8: SPE contribution plot sample instant 9619 showing the flowrate and compressor variables as the top contributors:Loss of Reaction

an increase in the compression ratio. For the reactor, this condition is not critical but it may cause an imbalance in the two stage secondary compressor leading to higher wear.

Fig. 6.10 shows the SPE remains at roughly the same value until sample 10075 after which there is a sharp rise. Fig.6.12 shows a bar plot of the contribution of all the variables at sample instant 10075. Similar to the contribution plots obtained in the discussion on decomposition in the Chapter 5, plots of contribution of all the 37 variables considered in the model were plotted in the manner described below,

- The SPE rises sharply after sample instant 10075 and Zone1B catalyst pump output becomes a significant contributor to the threshold violation.
- Percentage contribution of the variables was noted for every fifth sampling instant starting from sample 10115 to the final sample instant 10707.
- A combined plot of the contribution values of all the 37 variables was done and is shown in Fig. 6.14

The combined plot of all contributions for 121 observations gives us a picture of

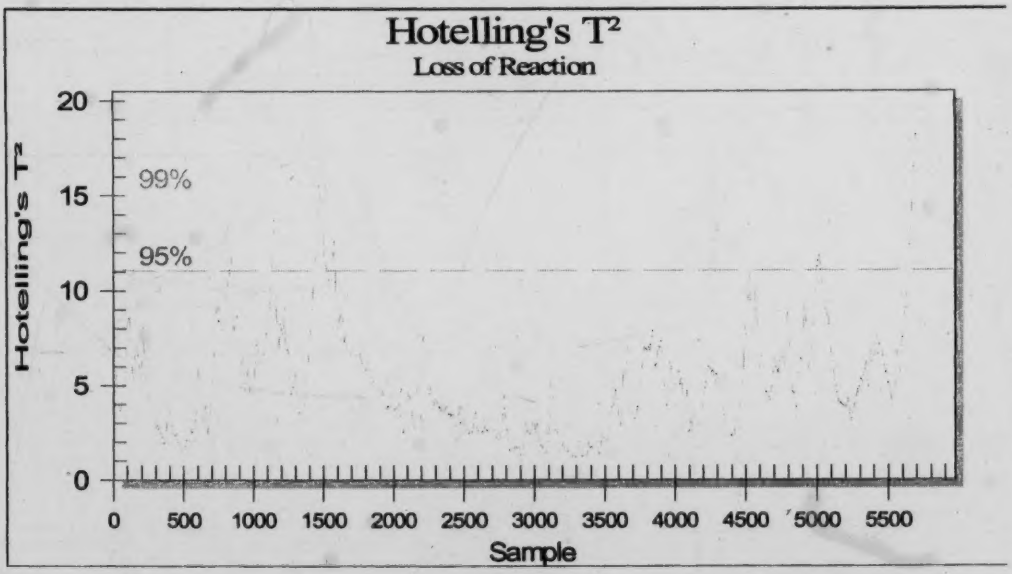


Figure 6.9: Hotelling's Plot showing violation of threshold:Loss of Reaction

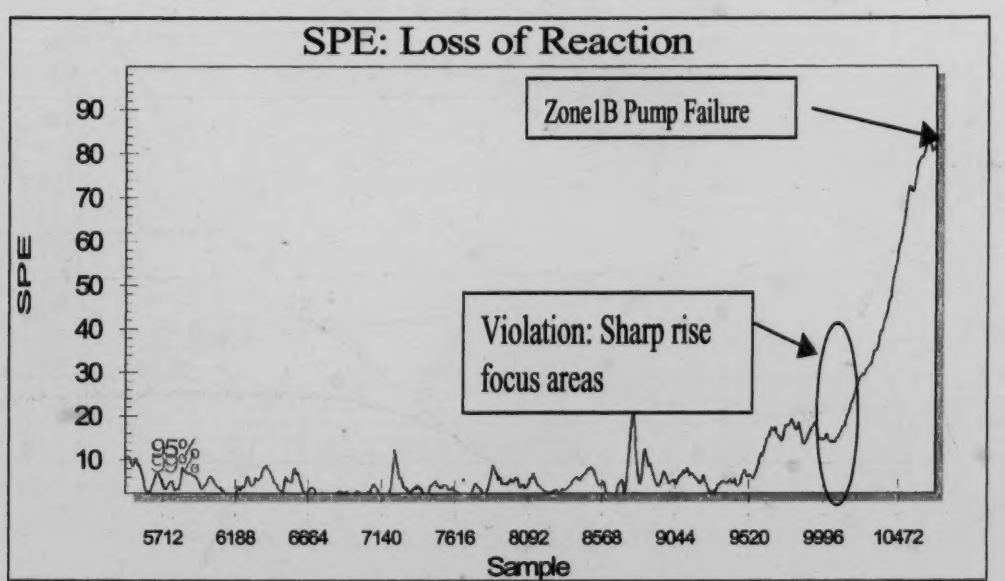


Figure 6.10: SPE:Loss of Reaction owing to catalyst pump failure

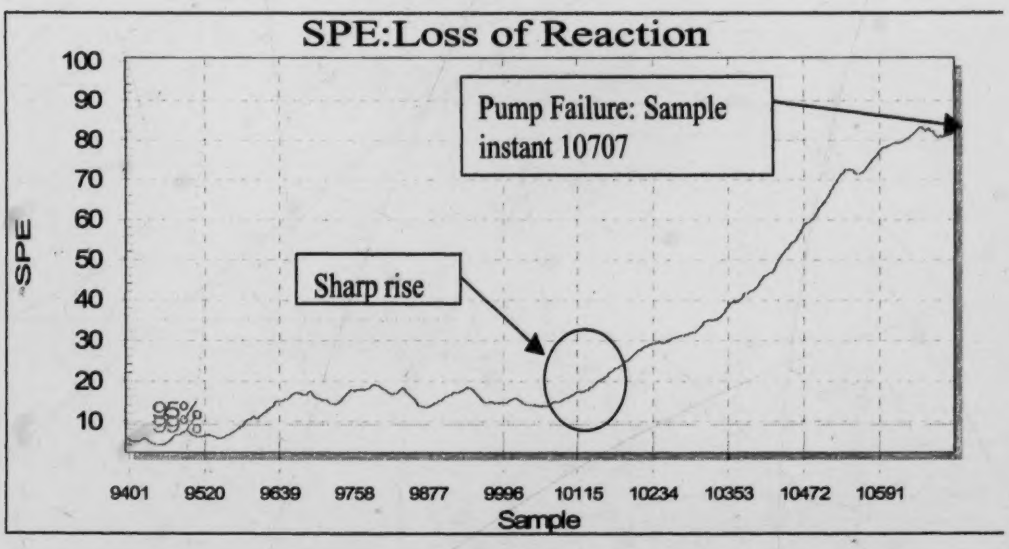


Figure 6.11: SPE:Loss of Reaction owing to catalyst pump failure

the event history leading to the fault(see Fig. 6.14). Flowrate , inter-stage pressure and secondary compressor second stage discharge temperature were initially the main contributors. All of these, however, combine to produce an undesirable variation in flowrate due to the variation in compressor suction temperature and interstage pressure. At sample instant 10420 (or observation 61) Zone1B catalyst pump output ($TIC52025D.OUT$) becomes a distinct contributor (over 10% contribution) and it remains a high contributor until the loss of reaction at sample instant 10707 (or observation 121). Meanwhile, the flowrate contribution is reduced to below 5% contribution. However, the interstage conditions remain high contributors throughout the time period studied.

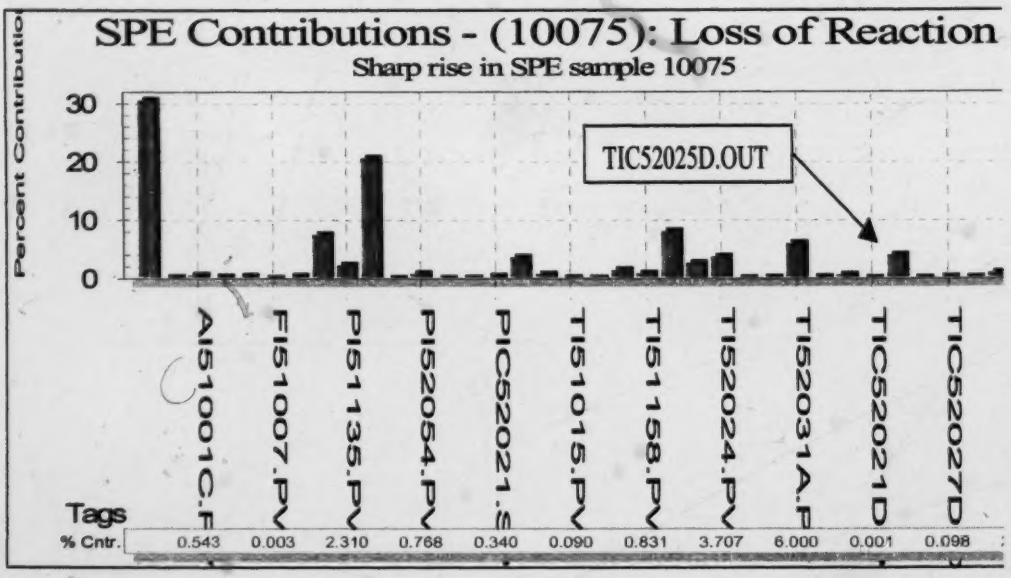


Figure 6.12: SPE contribution plot sample instant 10075 at which sharp rise starts in SPE starts:Loss of Reaction

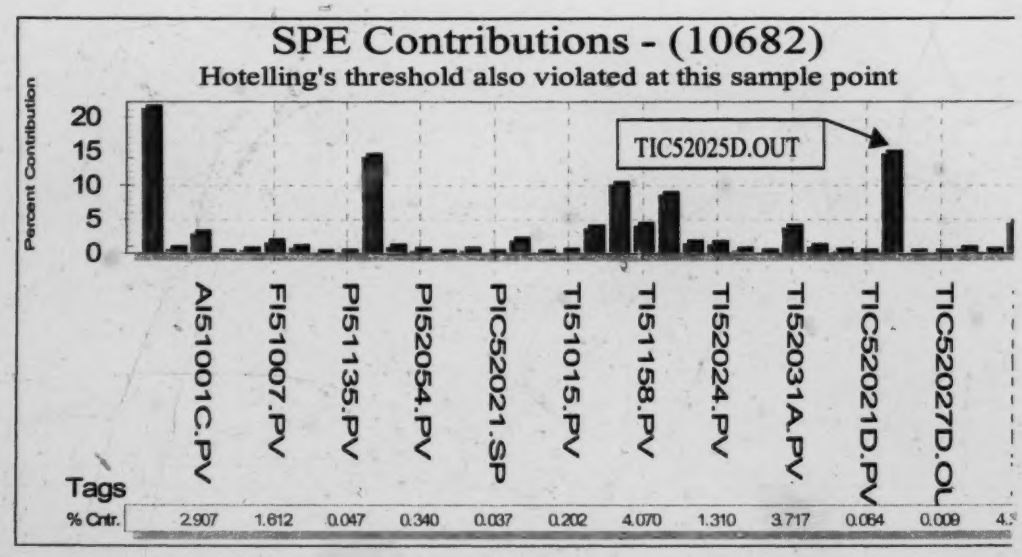


Figure 6.13: SPE contribution plot matching with sample instant at which Hotelling's threshold is also violated:Loss of Reaction

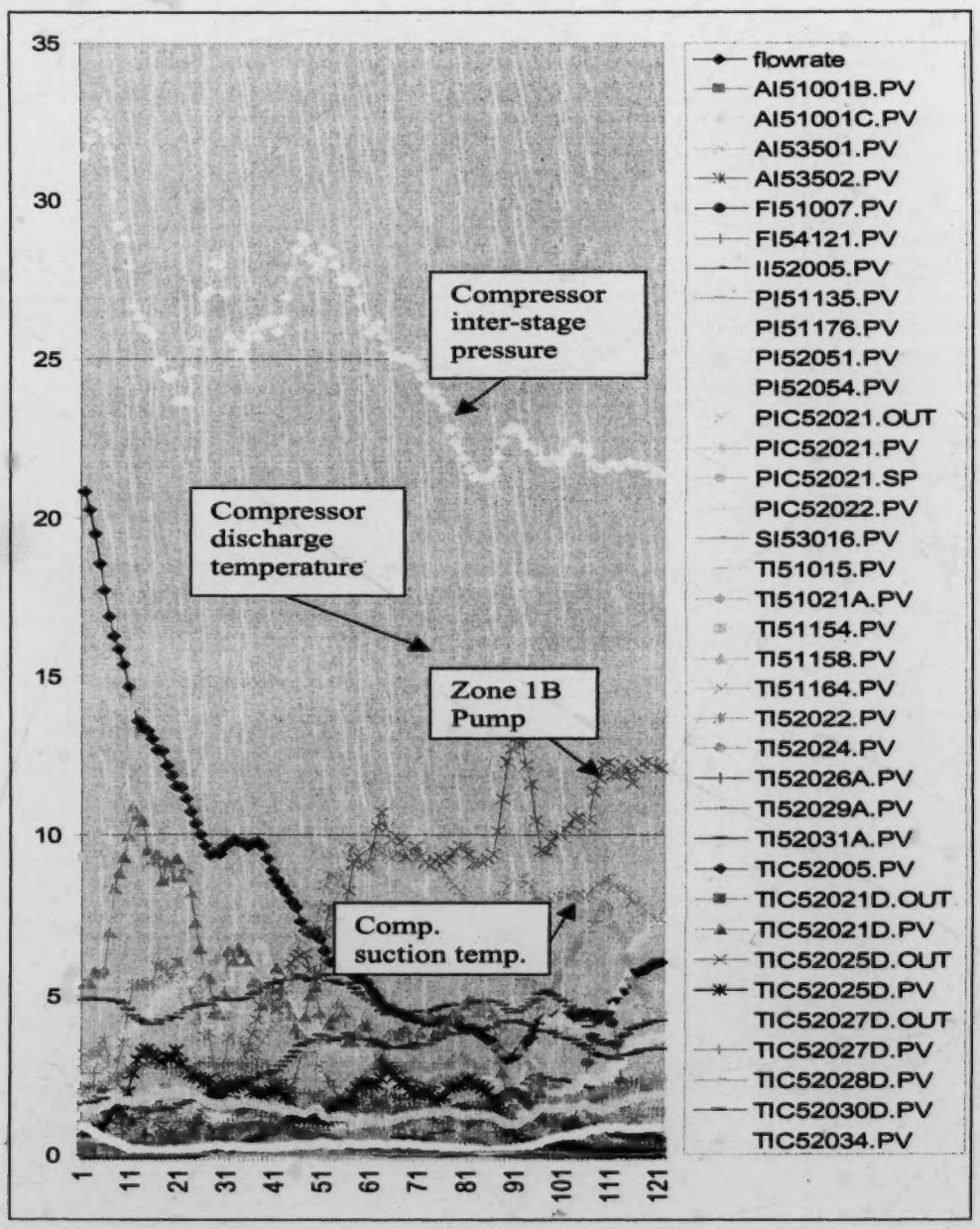


Figure 6.14: Trend Plot of Contribution of all variables in model:Loss of Reaction at monitoring point 121

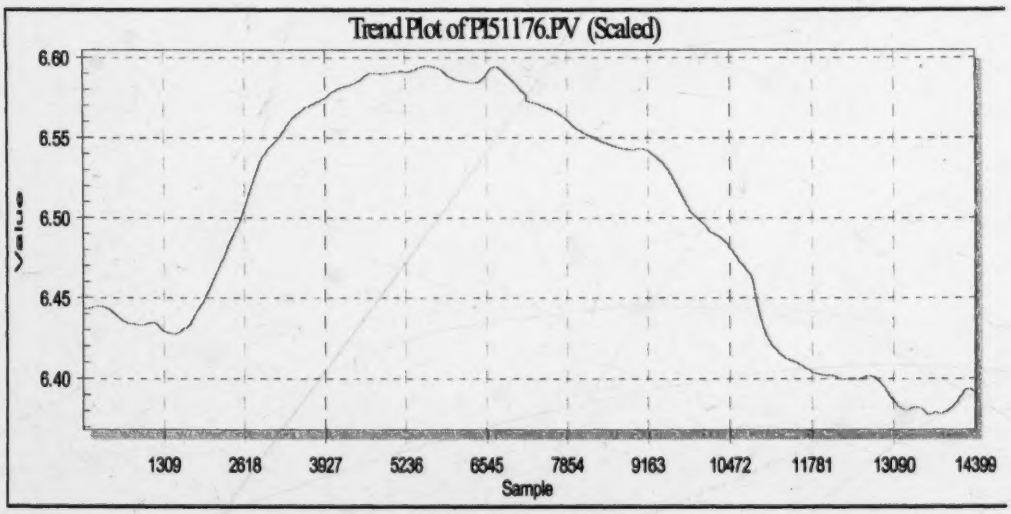


Figure 6.15: Trend Plot of interstage pressure PI51176.PV :Loss of Reaction

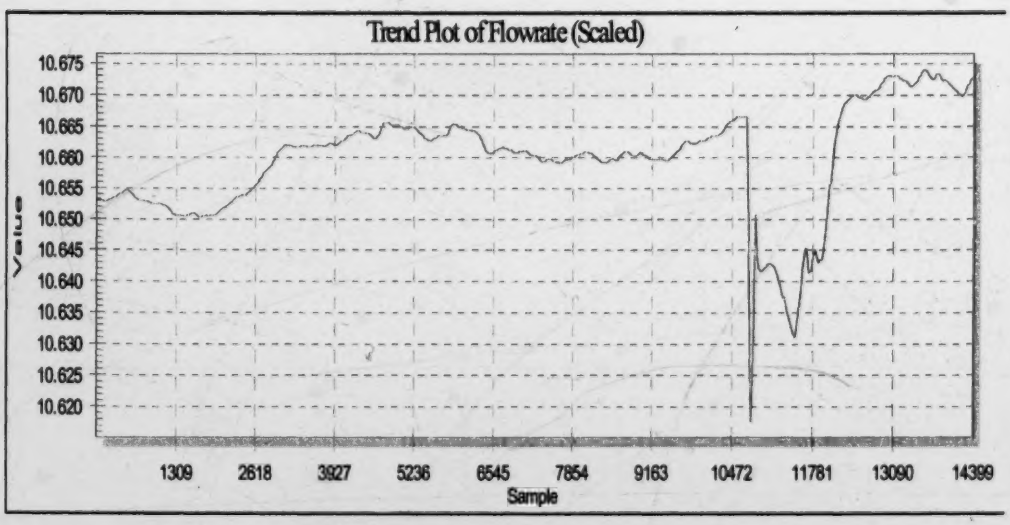


Figure 6.16: Trend Plot of Flowrate:Loss of Reaction

6.4 Conclusions

It is clear from the operation log and the PCA analysis that the reason for loss of reaction was the failure of the catalyst pump of Zone1B. Flowrate is a critical variable in the reaction kinetics but its contribution was not significant when the loss of reaction occurred. The conclusions of the analysis in this chapter are:

- Flowrate was initially an important contributor but fell to insignificant levels before the final loss of reaction.
- Interstage pressure and the associated final discharge temperature remained high contributors. Presently the cooling on the interstage is a single loop PID control. Since this is not a feed forward or multivariable control scheme, any change in the compressor suction conditions or the cooling fluid conditions will disturb the compressor variables until they are corrected by other actions. A redesigned control strategy for maintaining the interstage pressure under suction or cooling fluid disturbances needs to be devised. The compressor interstage and discharge variables are the effect of changes in suction conditions of the compressor. Using our process knowledge we see, there is clearly no connection in this case between the loss of reaction and the compressor variables
- Zone1B catalyst pump was changed over in a shift previous to the one in which the fault occurred. The first pump had run satisfactorily for more than four days of the same campaign.
- Apart from the compressor variables, the catalyst pump in Zone1B was the main contributor to the fault. **This is the most plausible reason for the loss of reaction.**

Chapter 7

Product Quality Design

7.1 Introduction

In the speciality polymer manufacturing business, various product specifications need to be satisfied for specific applications in diverse areas. The unique specification demands of the customer requires producers such as AT Plastics to continuously develop new polymer grades or products that are tailored to the required molecular structure to obtain the desired processing and physical properties. The quality design of new products is the job function of the technology researchers, chemists and engineers in industry and requires speed and efficiency to satisfy market demands and keep one step ahead of the competition.

Product quality design entails specifying 'recipes' or the set of operating conditions alongwith catalyst mix and property modifiers for manufacturing the target product. Analysis of reaction kinetics coupled with operating experience is utilized to specify the operating conditions. This technique is still very popular but has the following drawbacks:

- The relationship between the desired molecular structure and the kinetic model is not always well understood.
- Production trials need to be carried out which are costly since low quality off-grade material is made for most of the trials.
- Crucial time may be lost since trials generally require management approvals causing more delays and longer time to market.

A smarter, more efficient approach for product quality design was proposed by researchers at a polymer manufacturing facility at *Showa Denko, Japan* (Moteki and Arai 1988*a*) which combined multivariate statistical tools with reaction kinetic analysis for product quality design. Thereafter, little has appeared in the literature as extension to the quality design tools using their techniques.

The major advantage in using this approach is that it is a reliable tool for defining, based on real plant operation history, the set of process conditions required for a particular product quality. Thereafter, reaction kinetic analysis may be performed to determine possible operating conditions in terms of safety.

The objective of this chapter is to develop a simple tool based on multivariate statistical analysis which will define a set of process operating conditions that will result in the desired product quality. The set of process conditions may then be fine-tuned using reaction kinetics studies to obtain even higher product quality. This chapter proposes the steps to be followed for new product design by using multivariate statistical correlation charts and regressed property models.

7.2 Multivariate Statistical Tools for Product Quality Design

Development of multivariate statistical tools for generating 'recipes' or instructions for plant operation will give the engineers and research chemists a mathematical way to design new products based on real plant performance history. The steps to perform the product design are:

- Selection of on-line property and laboratory data for molecular structure and physical properties for all grades being manufactured in the plant. For the current work we have online melt index measurement (a macro physical property and an inverse measure of the viscosity and an inverse measure of Molecular weight). The other property measured online is the vinyl acetate (VA) comonomer content of the manufactured polymer. The VA content controls certain properties of the manufactured polymer such as its softening temperature. Other response quality variables which may be measured for future

work are T_m(melting point), T_c(Crystallization temperature), α(heat of fusion)and density. These properties will suffice to define the physical properties required for most applications. For special analyses we could also measure M_w(weight average molecular weight), M_n(number average molecular weight), MWD(molecular weight distribution), SCB(Short chain branching), LCB(Long chain branching), unsaturated groups, distribution of SCB and LCB, viscosity and ESCR (Environmental Stress Cracking resistance) (Moteki and Arai 1988*a*)

- Selection of critical operating parameters in the processing unit. This can be done using PCA analysis by rejecting variables with low loadings in the PC's in the model. For the current work, 36 plant operation variables and a calculated variable of mass flowrate have beep found to explain most of the variability in the plant.

- Collection of required laboratory data and operation data for the various grades manufactured in the plant.

- Development of regressed relationships between the product quality variables (response variables) and the operating data (predictor variables).

- Development of correlation charts for the quality variables and the plant operating data.

- Selection of a product closest to the desired properties.

- Determination of operating parameters required to obtain desired quality using correlation charts. The proposed process operating conditions are then introduced into the regressed relationships of product quality and optimized to obtain the final operating parameters. Finally, a check for operating feasibility is carried out using reaction kinetics and operating experience.

- Management approval and trials can then be carried out at the predicted conditions to ensure the desired product quality is achieved with the new set of operating conditions.

The most common use for LDPE and its copolymers is in film, lamination and adhesives and hence, the optical properties of haze, gloss, clarity etc. are important specifications. These depends on crystallinity, which in turn is related to the molecular structure with respect to SCB and LCB respectively. In addition to transparency, the melt processing characteristics and mechanical properties depend upon molecular weight and molecular weight distribution. Previous research (Moteki and Arai 1988*a*) shows that the dependant variable (product properties, which are also called response variables such as thermal properties, or molecular structures such as SCB and LCB) can be satisfactorily explained by regressed relationships using the plant operating conditions and the feed mix. However, the operating conditions required to obtain certain molecular weight and rheology are hardest to predict. With more powerful computing and alternative techniques using transformations, it may be possible to have higher accuracy regressed relationships between the response property variables and the process conditions.

In this work MI(melt index) and the Vinyl acetate content are available as online physical property measurements. Regressed relationships for MI and other quality variables have not been developed due to the non availability of exhaustive quality data and the focus during this study being on fault detection and diagnosis. The development of correlation matrices and simple approaches to developing greater insight to product quality design using PCA are explained in the subsequent sections. Correlation matrices for some grades appear in Appendix *E*.

7.3 Correlation studies

LDPE manufactured in the high pressure process is very commonly used for film and lamination. The properties of haze, gloss and clarity are important specifications which are highly correlated to crystallinity, SCB and LCB distributions and molecular weight distribution (Moteki and Arai 1988*a*). Therefore, if we are to use multivariate statistical tools for product quality design , we need to have reliable quantitative relationships like multivariate regressed models between structure (product quality or response variable) and the process operating conditions (process variables or predictor variables). It is also known from kinetic reaction analysis that the polymerization

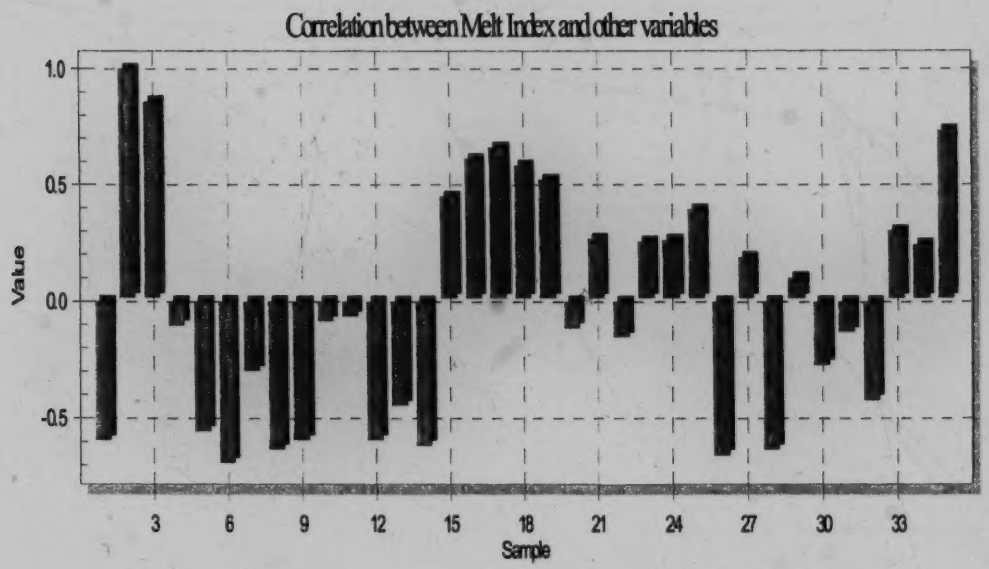


Figure 7.1: A Bar chart for MI correlation with plant parameters

conditions are highly correlated to each other and even small changes in pressure, temperature profile of reactor, feed temperature, mass flowrate, feed mix for copolymer, etc. will cause changes in both the physical properties and molecular structure of the polymer. Therefore, to apply quantitative relationships to design a product, we need to first have a thorough understanding of the correlation and its direction between both quality variables and process variables. The correlations between variables and their directions with respect to each other will guide the fixing of operating conditions required to obtain a desired product quality. Since the current work does not have exhaustive quality data, regressed relationships have not been developed. However, the techniques to apply the quantitative relationships for product quality design using correlation studies and PCA are explained in this section.

The correlation matrix for a product is developed by: collecting appropriate operating data with maximized excitation; preprocessing the data; auto scaling the data matrix and evaluating the covariance matrix. In other words the correlation matrix between process variables is in the form of a covariance matrix of the processed and scaled quality and operating data. A word of caution is that since the data is obtained under routine operating conditions, there may not be enough excitation in some in-

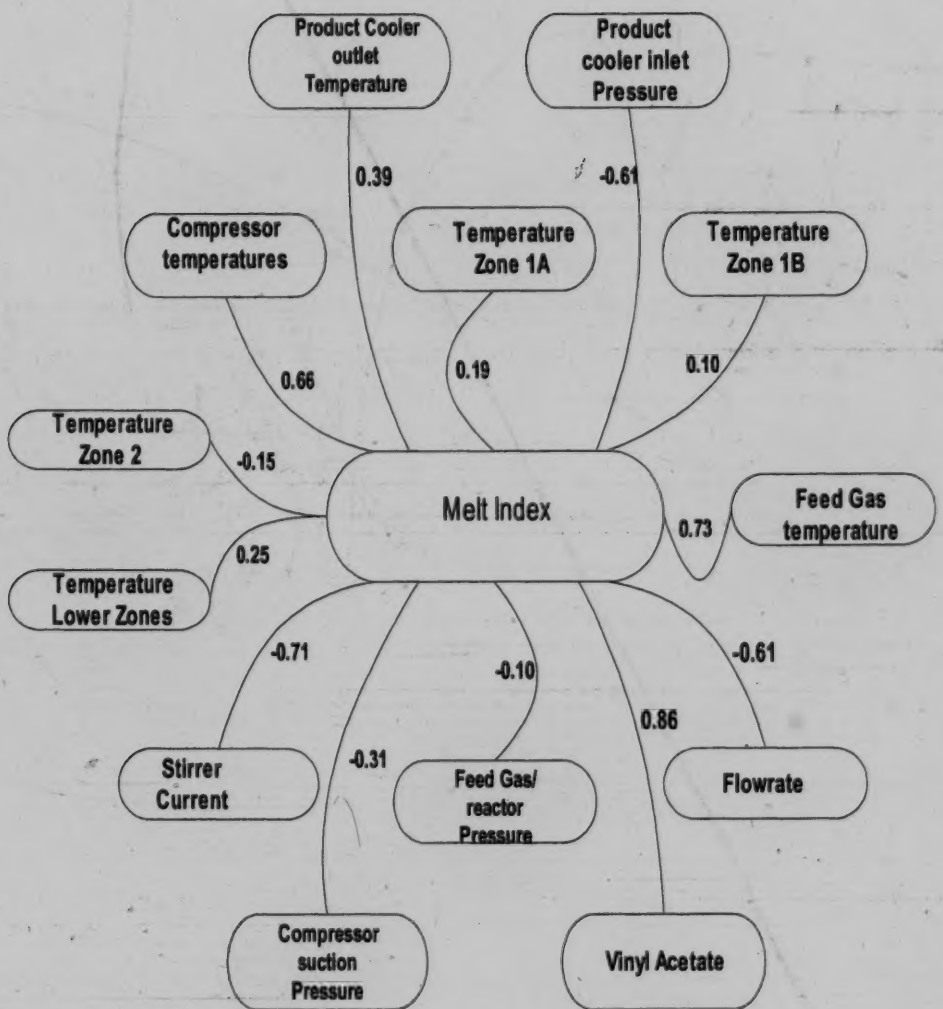


Figure 7.2: The organizational chart for correlation indices of MI with plant parameters.

put variables to see a strong correlation. Hence, some correlations which are obvious from process knowledge may not show up in the correlation matrix. To overcome this problem in part, the data has been chosen from cooler cook events together with normal operating data and the data matrix has been lagged using cross-correlation maximization and process knowledge before evaluation of the correlation matrix. *In future, the data matrix to be used for correlation studies should be chosen from normal operation which includes cooler cooks or grade slides and then lagged appropriately using cross-correlation maximization and process knowledge.*

Since melt index is the only property available for the current work the technique is used to design and control melt index of the product. In the case considered below, cooler cook and normal operating data from the 18% VA class of products is considered. Following the afore mentioned steps the correlation matrix was obtained. The observations were:

- See Fig. 7.1 for a bar chart of the correlation of the product's melt index with other variables of the plant. The same correlation is shown in another organizational chart (See Fig.7.2) giving the relative importance of variables affecting the melt index of the product. These plots provide a guide to the technology researchers, chemists and engineers in industry who are devising operating recipes for new products. The plots show that the melt index is negatively correlated to the mass flowrate. It is also negatively correlated to compressor pressure variables and positively correlated to the compressor temperature variables which are in effect variables guiding the flowrate through the system. Vinyl acetate content of the product is a strong factor for melt index control and shows up as highly positively correlated. Another interesting feature is the different directions in which the temperatures in the various zones of the reactor act on the melt index of the product. While the top zone is positively correlated to the the melt index, zone 2 is negatively correlated and the lower zones have positive correlations. The reactor pressure has negative correlation to melt index. Feed gas temperature, another important reaction variable has a fairly high positive correlation with melt index. This knowledge can go a long way in guiding recipe formation and operating guidelines as the critical features affecting the product

quality are highlighted. However, there is a word of caution and that is reaction kinetics explains that flowrate and melt index are positively correlated but the same is not seen in the correlation matrix we have obtained. One reason for this is that since feed gas temperature has a strong positive correlation to melt index and during cooler cooks the feed gas temperature is raised substantially which eclipses the normal effect of the product hardening (reduction in melt index) in the case of slowing down of the reactor (reduction in flowrate). In future data with the same feed gas temperature and varying flow rates could be used to study the effect of flowrate on melt index. Other anomalies like a low negative correlation between flowrate and the online vinyl acetate content measurement need to be studied in greater detail. The correlation between flowrate, vinyl acetate flow rate and the compressor variables is satisfactory in this analysis and follows our normal understanding of the process.

- Fig.7.3 is a color representation of the correlation matrix. It gives an overall picture of the correlations occurring amongst all the variables in the plant. Some features which are easily distinguishable are the strong correlations amongst the compressor variables, the reactor temperatures and the reactor catalyst pumps. Strong correlations exist amongst the top zone temperatures and the catalyst pumps. There is a correlation amongst the upper zone catalyst pumps injection rate with those in the lower zones and hence, the possibility of prediction for loss of reaction. The strong positive correlation between flowrate, the extruder speed, stirrer motor current and fresh feed rate follow our normal understanding of the process and give credence to the procedures adopted for fixing lags in the data matrix.

In addition to general guidelines to the direction a plant should take in controlling product quality, the above correlation plots can also be used in conjunction with regressed relationships for deciding on a recipe for a special product. The correlation charts give an overall view of the way in which a certain product quality is affected by various plant operating parameters . Small changes in operating variables can be made and inserted in to quantitative regressed relationships, between product quality and plant operating variables, to give the final operating conditions.

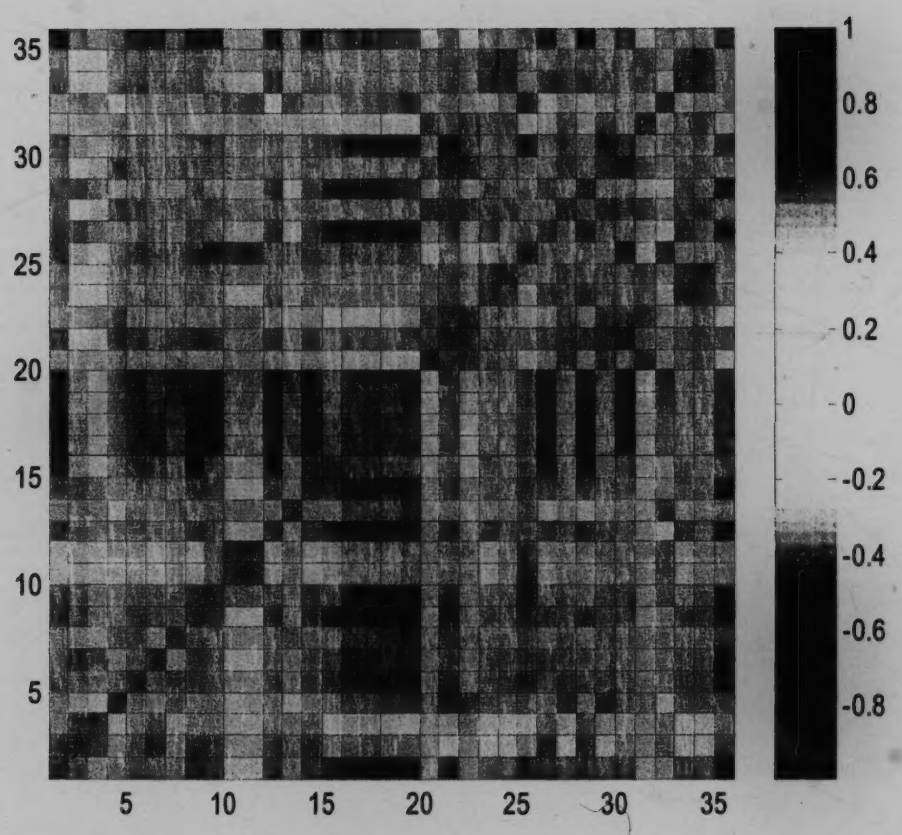


Figure 7.3: Correlation color matrix: Notice the correlations of parameters with MI (variable no.2), the highly correlated compressor variables (nos.15 to 19), the highly correlated reactor temperatures (nos.20 to 24), the highly correlated variables of Zone1 (nos.26 to 29).

7.4 PCA model: Information on dominant factors

The prime objective of PCA is to simplify and compress the operating and quality data into fewer factors which describe most of the variability of the plant. While selection of process variables can make a significant impact on PCA based analysis, it has been found that up to 90% of the character of the complicated process can be explained by five to seven PC's for all the grades of products manufactured and the special operations in the unit. The loadings of the various PC's give a macroscopic view of the correlations amongst the variables in a PC and the dominant factors controlling the quality of the polymer produced. The observations were as under,

- The first principal component, PC1 explains from 43% up to 52% of the variability in the process. This means that the process is controlled around 50% by the high loadings variables in the first PC. These are the flowrate, the reactor pressure and the product cooler outlet temperature in the positive direction and; the melt index, vinyl acetate content, stirrer motor current and the feed gas temperature in the negative direction.

- PC2, the second principal component covers most of the reactor temperatures. The dominant variables in this PC are the product cooler inlet pressure and zone 1 temperatures in the positive direction and; zone 2 and lower temperatures in the negative direction. The PC explains up to 20% of the plant variability.

7.5 Conclusions

The advantages in applying the above techniques for product quality design are:

- Production trials are minimized thus improving economic efficiency and plant safety.

- A reliable mathematical tool based on real plant performance history for product design can be used for improving product quality.

- The critical parameters for product quality control are highlighted. For example, if the top zone of the reactor plays the greatest role in defining the product

quality for a certain class of products, the industry move to selectively have better control strategies for the top zone temperatures.

- There will be an increase in confidence to quote for new 'grades' of product for customers. The structure and properties of speciality polymers from competing industry can be measured and quality design tools based on multivariate statistical analysis can be used be to seek operating conditions to develop the same quality of polymers.

Chapter 8

Conclusions

8.1 Contributions of this Thesis

In summary the contributions of this thesis are:

- **It has been shown using real industrial data from a high pressure polymerization unit that PCA can be successfully employed for fault detection and diagnosis.**

  Widely researched issues like decomposition of reactants can be detected as gross deviations from the 'normal' operation regime. Loss of reaction in reactors due to catalyst pump failures can also be predicted since the operating variables in the reactor are highly correlated. Using a combination of contribution plots and process knowledge, on-line schemes may be devised to avoid recurrence of such events.

- **Only similar operating events can be compared with each other.**

  It has been established that the model set up for analyzing an event has to be from similar operating conditions with similar variance information. The decomposition fault detection and diagnosis could be done only after a model was set up using a series of cooler cook operating data appended together from a similar class of 18% VA content products. Additionally, spare pumps may be able to deliver the same flow rate at different loadings and may generate a bias in the on-line monitoring index in case of change over(refer §Section 6.2). This requires mean correction of data in case the spare equipment is taken in line during the monitoring of the same campaign.

- **A first principles model has been made for the compressor to calculate the mass flow rate of reactants through the reactor.**

  The model has an ability to evaluate flow rate by taking an observation vector as input at every sampling instant and can be implemented as an on-line algorithm. The flowrate evaluation has helped the resolution of the factors actually responsible for the faults as it is established that flowrate dips or turndown of up to 66% of the rated flowrate can be be handled by the process without the occurrence of a fault.

- **The chapter on product quality design lays out simple and effective ways to improve 'recipe' formation and process monitoring.** The correlation charts give a mathematical handle to the technology researchers, chemists and engineers in industry to develop new products or 'grades' in combination with quantitative regressed relationships. There is however, a word of caution and that is that regressed relationships may have error if used outside the range of observations and the same has been experienced internally by *ICI* researchers. One way to work around this problem is to choose a grade already being manufactured closest to the one being developed and then check or compare the suggested new operating conditions for the target grade using the regressed relationships. The new operating data can then be cleared from a process feasibility and safety angle by using the knowledge of reaction kinetics and operating experience. The correlation charts also point to the critical variables which may need advanced strategies for better control to improve product quality. It is also established that cooler cook or grade slide data should be used to develop correlation charts and for future implementation of model predictive or multivariable control strategies, as this data has the maximum excitation.

8.2 Recommendations for Future Work

Need for Online Recursive PCA: PCA models built from historical data are time-invariant while the process plants are time-varying. This is a limitation in case online implementation of the scheme is to be carried out to monitor the process using

the monitoring indices of Hotelling's $\mathbf{T^2}$ and SPE. It has been seen in this study that the PCA models may change due to:

- Changes in operating means and variances which occur with change in plant throughput and equipment.
- Changes in the correlation matrix and the number of PC's in the model due to process unit revamps or equipment replacement.

In order to minimize false alarms, the need to recursively update models and the threshold limits for the monitoring indices is required. This will also give a handle to operators since they may re-initialize the monitoring system in case of special operations and equipment changes. In the case of grade slide and the initiation of a new campaign the system may be re-initialized to set up a new operating model which can be retained for the rest of the campaign for monitoring until there is an equipment change or a special operating event.

Here, we list some directions for future research:

- The major hurdle of establishing the fact that PCA can be used to detect many major faults using industrial data has been crossed. To derive real benefits the scheme has to be implemented on-line using recursive techniques. The issues of on line data access followed by adequate data pre-processing and evaluation of flowrate in a time efficient manner have to be researched and implemented. Recent research on this area has been reported in (Li and Qin 2000).
- Post on-line implementation, operator training has to be carried out for diagnosis and corrective action in case a fault is detected by the monitoring indices. This step calls for the development of user friendly interaction methods whereby the operator can effectively call for contribution plots and then zero-in using variable elimination methods based on memory and experience. Prior to implementation of the scheme it is strongly recommended to automate cooler cooks and grade slides. Since these are routine and very frequent operations, automation will standardize these events thereby improving the resolution of real process faults.

- The chapter on product quality design is an initiation of product design by a mathematical method. This can become a crucial and important tool in real product design. Future research should be geared towards developing quantitative relationships for all the important quality variables and setting up correlation charts for all grades.

Bibliography

Bett, K. E. (2000). *Kirk-Othmer Encyclopedia of Chemical Technology.*

Britton, L.G., Taylor D.A. and D.C. Wobser (1986). Thermal stability of ethylene at elevated pressures. *Plant/Operations Progress* **5**(4), 238–251.

Chiang, L.H., Russel E. L. and R. D. Braatz (2000). *Fault Detection and Diagnosis in Industrial Systems.* Advanced textbooks in control and signal processing. Springer.

Himes, D. M., Storer R. H. and C. Georgakis (1994). Determination of the number of principal components for disturbance detection and isolation. *In Proc. Of the American Control Conf.* pp. 1279–1283.

Hotelling, H. (1947). *Techniques of Statistical Analysis.* Chap. Multivariate quality control. McGraw Hill. New York.

Jackson, J. E. (1991). *A User'S Guide To Principal Components.* John Wiley & Sons. New York.

Kanury, A M (1977). *Introduction to Combustion Phenomena.* Gordon and Breach. New York.

Kasper, M H and W H Ray (1992). Chemometric methods for process monitoring and high-performance controller design.. *AIChE J.* **38**, 1593–1608.

Kim, J. and M. Larsen (1998). Improving quality with integrated statistical tools. *ASQ's 52nd Annual Quality Congress Proceedings* pp. 218–225.

Kosanovich, K A, Piovoso M J Dahl K S MacGregor J F and P Nomikos (1994). Multi-way pca applied to an industrial batch process.. pp. 1294–1298.

Kourti, T. and J. F. MacGregor (1998). Projection methods for process analysis and statistical process control: Examples and experiences from industrial applications. *ASQ'S 52nd Annual Quality Proceedings* pp. 73–77.

Kresta, J. V., Marlin T. E. and J. F. MacGregor (1991). Multivariable statistical monitoring of process operating performance.. *Can. J. of Chem. Eng.*, **69**, 35–47.

Ku, W., Storer R. H. and C. Georgakis (1995). Disturbance detection and isolation by dynamic principal component analysis.. *Chemometrics and Intelligent Laboratory Systems* **30**, 179–196.

Li, W., Yue H. Valle-Cervantes S. and S.J. Qin (2000). Recursive pca for adaptive process monitoring. *Journal of Process Control* **10**, 471–486.

Luo, R., Misra M. and D. M. Himmelblau (1999). Sensor fault detection via multiscale analysis and dynamic pca.. *Ind. Eng. Chem. Res.* **38**, 1489–1495.

MacGregor, J F (1994). Statistical process control of multivariate processes. *In proc. Of the IFAC Int. Symp. on Advanced Control of Chemical Processes* pp. 427–435.

MacGregor, J. F. and T. Kourti (1995). Statistical process control of multivariate processes. *Control Engineering Practice* **3**, 403–414.

MacGregor, J. F., Skagerberg B. and C. Kiparissides (1991). Multivariate statistical process control andproperty inference applied to low density polyethylene reactors. *IFAC symposium series, ADCHEM'91* (8), 155–159.

Mardia, K. V., Kent J. T. and J. M. Bibby (1982). *Multivariate Analysis.* Academic Press. London.

Miller, P., Swanson P. E. and C. E. Heckler (1998). Contribution plots:a missing link in multivariate quality control. *Applied Mathematics and Computer Science* **8**, 775–792.

Moteki, Y. and Y. Arai (1988*a*). Operation planning and quality design of a polymer process. *IFAC Proceedings Series* (4), 159–165.

Moteki, Y. and Y. Arai (1988*b*). Operation planning and quality design of a polymer process. *IFAC proceedings series* (4), 159–165.

Pearson, K (1901). On lines and planes of closest fit to systems of points in space. *Philos. Mag.* **2**, 559–572.

Petrovich, M. V. (1998). Performance analysis for process improvement. *ASQ'S 52nd Annual Quality Proceedings* pp. 697–707.

Piovoso, M J and K A Kosanovich (1994). Applications of multivariate statistical methods to process monitoring and controller design.. *Int. J. of Control* **59**, 743–765.

Piovoso, M J, Kosanovich K A and R K Pearson (1992). Monitoring process performance in real time.. pp. 2359–2363.

Raich, A C and A Cinar (1994). Multivariate statistical methods for monitoring continuous processes: Assessment of discriminatory power disturbance models and diagnoses of multiple disturbances.. *Chemometrics and Intelligent Laboratory Systems* **30**, 37–48.

Raich, A. C. and A. Cinar (1995). Statistical process monitoring and disturbance diagnoses in multivariable continuous processes. *AIChE J.* **42**, 995–1009.

Seborg, D. E., Edgar T. F. and D. A. Mellichamp (1989). *Process Dynamics and Control.* John Wiley and Sons. New York.

Skagerberg, B., Lehtenin J. and Wiklund Jan (1992). Multivariate process monitoring of polypropylene production. *Plastics Rubber and Composites Processing and Applications* **18**(5), 299–305.

van Heerden, C. (1953). Autothermic processes.properties and reactor design. *Industrial and engineering chemistry* **45**(6), 1242–1247.

van Heerden, C. (1958). The character of the stsionary state of the exothermic processes. *Chemical Engineering Science* **8**, 133–145.

Wise, B M and N B Gallagher (1996). The process chemometrics approach to process monitoring and fault detection.. *J. of Process Control* **6**, 329–348.

Wise, B. M., Ricker N. L. Veltkamp D. F. and B. R. Kowalski (1990). A theoretical basis for the use of principal component models for monitoring multivariate processes. *Process and Control Quality* **1**, 41–51.

Wold, S., Esbenson K. and P. Geladi (1987). Principal components analysis. *Chemometrics and Intelligent Laboratory Systems* p. 2:37.

Zhang, Haitao (2000). Statistical process monitoring and modelling using pca and pls. Master's thesis. Alberta.

Zheng, L. L., McAvoy T. J. Huang Y. and Gang Chen (2001). Application of multivariate analysis in batch processes. *Ind. Eng. Chem. Res.* **40**, 1641–1649.

Appendix A

Process Variables Logged by Data Server

AT Plastics Inc.		
1	AI51001B.PV	Gas A (PV)
2	AI51001C.PV	Gas B (PV)
3	AI51001D.PV	Gas C (PV)
4	AI53501.PV	M.I (PV).
5	AI53502.PV	V.A. (PV)
6	FI51007.PV	Fresh Feed Flow Rate (PV)
7	FI54010.PV	Modifier Mass Flow (PV)
8	FI54121.PV	Total Comonomer Mass Flow (PV)
9	II52005.PV	Stirrer Motor Amps (PV)
10	LIC52001.PV	Seperator Level (PV)
11	PI51164.PV	Compressor Pressure P1 (PV)
12	PI51176.PV	Compresser Pressure P2 (PV)
13	PI52051.PV	Product Cooler Presure P3 (PV)
14	PI52054.PV	Reactor Pressure P4 (PV)
15	PIC52021.OUT	Reactor Pressure P5 (OUT)
16	PIC52021.PV	Reactor Pressure P5 (PV)
17	PIC52021.SP	Reactor Pressure P5 (SP)
18	PIC52022.PV	Separator Pressure P6 (PV)
19	SI53016.PV	Extruder Speed (PV)
20	TI51015.PV	Return Gas Cooler Temp T1 (PV)
21	TI51021A.PV	Return Gas Cooler Temp T2 (PV)
22	TI51154.PV	Compressor Temp T3 (PV)
23	TI51158.PV	Compressor Temp T4 (PV)
24	TI51164.PV	Compressor Temp T5 (PV)
25	TI52022.PV	Reactor Thermo Couple TR1 (PV)
26	TI52024.PV	Reactor Thermo Couple TR2 (PV)
27	TI52026A.PV	Reactor Thermo Couple TR3 (PV)
28	TI52029A.PV	Reactor Thermo Couple TR4 (PV)
29	TI52031A.PV	Reactor Thermo Couple TR5 (PV)
30	TIC52005.PV	Product Cooler Temp TP1 (PV)
31	TIC52021D.OUT	Cat. Injection Pump TC1D (Output)
32	TIC52021D.PV	Reactor Thermo Couple TC1D (PV)
33	TIC52021D.SP	Reactor Thermo Couple TC1D (Set Point)
34	TIC52021E.OUT	Reactor Thermo Couple TC1E (Output)
35	TIC52025D.OUT	Cat. Injection Pump TC2D (Output)
36	TIC52025D.PV	Reactor Thermo Couple TC2D (PV)
37	TIC52025D.SP	Reactor Thermo Couple TC2D (Set Point)
38	TIC52025E.OUT	Cat. Injection Pump TC2E (Output)
39	TIC52027D.OUT	Cat. Injection Pump TC3D (Output)
40	TIC52027D.PV	Reactor Thermo Couple TC3D (PV)
41	TIC52027D.SP	Reactor Thermo Couple TC3D (Set Point)
42	TIC52027E.OUT	Cat. Injection Pump TC3E (Output)
43	TIC52028D.OUT	Cat. Injection Pump TC4D (Output)
44	TIC52028D.PV	Reactor Thermo Couple TC4D (PV)
45	TIC52028D.SP	Reactor Thermo Couple TC4D (Set Point)
46	TIC52028E.OUT	Cat. Injection Pump TC4E (Output)
47	TIC52030D.PV	Reactor Thermo Couple TC5D (Set Point)
48	TIC52034.PV	Feed Gas Temp TF1 (PV)

Figure A.1: Process Variables logged by server

Appendix B

Raw Data File

Tagname: PIC52021.PV

ip_trend_time	ip_trend_value	ip_trend_qstatus
07-FEB-02 18:00:00.2	1399.900	Good
07-FEB-02 18:00:01.3	1399.900	Good
07-FEB-02 18:00:02.3	1399.600	Good
07-FEB-02 18:00:03.3	1399.600	Good
07-FEB-02 18:00:04.2	1399.600	Good
07-FEB-02 18:00:05.2	1399.600	Good
07-FEB-02 18:00:06.2	1399.600	Good
07-FEB-02 18:00:07.2	1399.600	Good
07-FEB-02 18:00:08.2	1399.600	Good
07-FEB-02 18:00:09.2	1399.600	Good
07-FEB-02 18:00:10.3	1399.600	Good
07-FEB-02 18:00:11.3	1399.600	Good
07-FEB-02 18:00:12.3	1399.600	Good
07-FEB-02 18:00:13.3	1399.600	Good
07-FEB-02 18:00:14.2	1399.900	Good
07-FEB-02 18:00:15.3	1399.900	Good
07-FEB-02 18:00:16.3	1399.900	Good
07-FEB-02 18:00:17.2	1399.900	Good
07-FEB-02 18:00:18.2	1399.900	Good
07-FEB-02 18:00:19.3	1399.900	Good
07-FEB-02 18:00:20.3	1400.300	Good
07-FEB-02 18:00:21.3	1400.300	Good
07-FEB-02 18:00:22.2	1400.300	Good
07-FEB-02 18:00:23.3	1400.300	Good
07-FEB-02 18:00:24.2	1400.300	Good
07-FEB-02 18:00:25.2	1400.300	Good
07-FEB-02 18:00:26.2	1399.900	Good
07-FEB-02 18:00:27.3	1399.900	Good
07-FEB-02 18:00:28.3	1399.900	Good
07-FEB-02 18:00:29.3	1399.900	Good
07-FEB-02 18:00:30.2	1399.900	Good
07-FEB-02 18:00:31.2	1399.900	Good
07-FEB-02 18:00:32.2	1400.600	Good
07-FEB-02 18:00:33.2	1400.600	Good
07-FEB-02 18:00:34.2	1400.600	Good
07-FEB-02 18:00:35.2	1400.600	Good

Figure B.1: Raw Data File Format

Appendix C

Cooler Cook Model Loadings

Variable	PC-1	PC-2	PC-3	PC-4
flowrate	0.92	-0.31	0.19	-0.02
AI51001Bpv	-0.23	0.59	-0.37	-0.62
AI51001Cpv	-0.63	-0.26	0.03	0.6
FI54121pv	0.82	0.01	-0.36	0.08
II52005pv	0.81	-0.02	0.09	-0.23
PI51135/64pv	0.86	-0.23	-0.16	0.26
PI51176pv	0.83	-0.38	0.29	-0.22
PI52051pv	0.67	-0.6	0.23	0.15
PI52054pv RFG Pr	-0.23	-0.94	-0.14	-0.04
PIC52021sp	-0.31	-0.91	-0.09	-0.17
PIC52021pv	-0.25	-0.92	-0.21	-0.1
PIC52021out	0.85	0.36	0.28	0.13
PIC52022pv	0.3	-0.58	0.34	0.3
SI53016pv	0.9	-0.03	0.21	0.04
TI51015pv	-0.86	0.19	0.09	0.28
TI51021Apv	-0.94	0.25	-0.05	-0.01
TI51164pv 1	-0.93	0.26	-0.16	0.01
TI51158pv 1	-0.95	0.2	0.01	-0.04
TI51154pv 2	-0.95	0.16	0.06	-0.11
TI52022pv 1A	-0.13	-0.1	0.94	-0.08
TI52024pv	-0.76	0.07	0.61	-0.1
TI52026Apv	-0.47	-0.08	0.85	-0.14
TI52029Apv	0.31	0.93	0.09	0.08
TI52031Apv	0.4	0.9	0.08	0.05
TIC52005pv	-0.86	0.07	0.12	0.34
TIC52021Dout	0.85	-0.38	0.27	-0.1
TIC52021Dpv .	-0.6	0.05	0.73	-0.04
TIC52025Dout	0.78	-0.49	0.32	0.02
TIC52025Dpv	-0.71	0.05	0.67	-0.1
TIC52027Dout	0.84	0.48	-0.06	0.15
TIC52027Dpv	0.39	0.84	0.26	0.13
TIC52028Dout	0.65	0.63	0.15	-0.13
TIC52028Dpv	0.26	0.95	0.09	0.07
TIC52030Dpv	0.4	0.9	0.08	0.03
TIC52034pv RFG T	-0.87	0.26	-0.34	0.07

Figure C.1: Cooler Cook Model Loadings

Appendix D

Loss of Reaction Model Variables and Loadings

Variable	PC-1	PC-2	PC-3	PC-4	PC-5
flowrate	0.91	-0.25	0.08	-0.06	0.13
AI51001B.PV	0.94	-0.07	-0.16	0.13	0.06
AI51001C.PV	0.09	-0.39	-0.09	0.53	0.42
AI53501.PV	-0.97	0.10	-0.01	-0.03	-0.15
AI53502.PV	-0.78	-0.36	-0.06	0.15	0.25
FI51007.PV	0.37	-0.19	-0.15	0.25	-0.23
FI54121.PV	0.65	-0.21	-0.38	0.34	0.30
II52005.PV	-0.69	0.16	0.10	-0.49	-0.15
PI51135.PV	-0.83	0.25	0.24	-0.19	-0.28
PI51176.PV	-0.92	0.23	0.18	-0.06	-0.19
PI52051.PV	-0.62	0.59	-0.08	0.42	-0.02
PI52054.PV	0.93	0.11	0.12	-0.18	-0.10
PIC52021.OUT	-0.83	-0.51	-0.02	0.12	-0.07
PIC52021.PV	0.95	0.23	0.15	-0.03	-0.03
PIC52021.SP	0.95	0.26	0.10	-0.08	0.02
PIC52022.PV	0.92	0.20	0.25	0.09	-0.04
SI53016.PV	0.06	-0.19	0.05	0.26	-0.47
TI51015.PV	0.94	0.26	0.06	0.02	-0.05
TI51021A.PV	-0.15	-0.02	-0.43	-0.65	0.46
TI51154.PV	0.41	-0.38	-0.56	-0.30	0.29
TI51158.PV	-0.60	-0.57	0.00	-0.42	0.21
TI51164.PV	0.66	0.54	0.02	-0.19	0.13
TI52022.PV	0.39	0.48	-0.65	-0.09	-0.09
TI52024.PV	-0.52	0.55	-0.59	0.03	-0.15
TI52026A.PV	0.01	0.67	-0.62	-0.01	-0.28
TI52029A.PV	0.41	-0.84	-0.23	-0.07	-0.24
TI52031A.PV	0.21	-0.90	-0.07	-0.11	-0.31
TIC52005.PV	0.86	0.45	0.05	-0.01	-0.10
TIC52021D.OUT	-0.90	-0.36	-0.09	-0.01	0.04
TIC52021D.PV	0.00	0.47	-0.77	-0.05	-0.08
TIC52025D.OUT	-0.32	-0.60	-0.51	0.25	0.23
TIC52025D.PV	-0.16	0.68	-0.66	0.04	-0.20
TIC52027D.OUT	-0.13	-0.96	-0.14	0.00	-0.12
TIC52027D.PV	0.51	-0.72	-0.36	-0.02	-0.24
TIC52028D.PV	0.52	-0.79	-0.20	-0.05	-0.23
TIC52030D.PV	0.32	-0.85	-0.12	-0.14	-0.34
TIC52034.PV	-0.90	-0.05	-0.24	0.24	0.04

Figure D.1: Loss of Reaction model variables and loadings

Appendix E

Correlation Matrix for 18% VA Product Cooler Cook

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	
Flowrate	1	1	-0.61	-0.3	0.54	0.79	0.74	0.64	0.99	0.89	0.34	0.32	0.93	0.52	0.87	-0.8	-0.99	-0.99	-0.97	-0.91	-0.19	-0.66	-0.26	-0.5	-0.49	-0.68	0.89	-0.54	0.82	-0.55	0.8	-0.98	0.44	-0.58	-0.48	-0.94	1
AI52001.PV	2	-0.61	1	0.88	-0.12	-0.57	-0.71	-0.31	-0.65	-0.61	-0.1	-0.08	-0.61	-0.48	-0.83	0.45	0.61	0.66	0.58	0.52	-0.13	0.27	-0.17	0.26	0.28	0.4	-0.67	0.19	-0.65	0.1	-0.29	-0.15	-0.44	0.3	0.25	0.73	2
AI52002.PV	3	-0.3	0.88	1	0.08	-0.31	-0.57	-0.03	-0.35	-0.32	0.23	0.05	-0.35	-0.28	-0.36	0.23	0.33	0.37	0.29	0.21	-0.18	0.04	-0.3	0.19	0.21	0.25	-0.45	0.04	-0.47	-0.1	-0.03	-0.16	-0.32	0.19	0.18	0.5	3
FI51007.PV	4	0.54	-0.12	0.08	1	0.61	0.39	0.63	0.53	0.5	0.34	0.36	0.51	0.18	0.4	-0.41	-0.54	-0.51	-0.59	-0.6	-0.48	-0.71	-0.59	-0.45	-0.45	-0.46	0.47	-0.62	0.4	-0.09	0.61	-0.4	0.21	-0.5	-0.44	-0.46	4
FI54121.PV	5	0.79	-0.57	-0.31	0.61	1	0.69	0.78	0.82	0.78	0.4	0.4	0.77	0.53	0.99	-0.99	-0.8	-0.8	-0.84	-0.86	-0.33	-0.71	-0.35	-0.52	-0.57	-0.71	0.72	-0.82	0.88	-0.81	0.88	-0.25	0.49	-0.6	-0.58	-0.8	5
II52006.PV	6	0.74	-0.71	-0.57	0.39	0.69	1	0.47	0.77	0.73	0.21	0.29	0.71	0.43	0.98	-0.84	-0.76	-0.77	-0.74	-0.69	-0.21	-0.55	-0.24	-0.5	-0.53	-0.61	0.76	-0.53	0.84	-0.47	0.55	-0.23	0.51	-0.58	-0.51	-0.84	6
PI51135.PV	7	0.64	-0.31	-0.03	0.83	0.79	0.47	1	0.69	0.66	0.43	0.42	0.83	0.73	0.92	-0.45	-0.64	-0.62	-0.72	-0.83	-0.35	-0.62	-0.38	-0.4	-0.46	-0.55	0.5	-0.59	0.49	-0.58	0.82	-0.3	0.52	-0.47	-0.44	-0.54	7
PI51176.PV	8	0.99	-0.65	-0.35	0.53	0.82	0.77	0.68	1	0.86	0.33	0.31	0.94	0.54	0.97	-0.81	-0.99	-0.98	-0.97	-0.92	-0.19	-0.89	-0.35	-0.52	-0.5	-0.7	0.9	-0.53	0.84	-0.58	0.78	-0.07	0.45	-0.6	-0.5	-0.95	8
PI52051.PV	9	0.89	-0.61	-0.32	0.5	0.75	0.73	0.65	0.89	1	0.66	0.65	0.71	0.85	0.78	-0.54	-0.9	-0.9	-0.91	-0.89	-0.2	-0.66	-0.33	-0.33	-0.5	-0.76	0.85	-0.54	0.85	-0.58	0.81	-0.13	0.71	-0.42	-0.49	-0.81	9
PI52054.PV	10	0.34	-0.1	0.03	0.34	0.4	0.31	0.43	0.33	0.66	1	0.66	0.06	0.52	0.38	-0.05	-0.36	-0.37	-0.46	-0.49	-0.23	-0.37	-0.35	0.08	-0.3	-0.68	-0.09	-0.33	-0.12	-0.39	0.57	-0.19	0.75	0	-0.27	-0.25	10
PIC52021.PV	11	0.32	-0.08	0.05	0.36	0.4	0.29	0.42	0.31	0.65	0.66	1	0.04	0.48	0.25	0	-0.36	-0.34	-0.43	-0.47	-0.26	-0.36	-0.39	0.09	-0.3	-0.65	-0.09	-0.36	-0.16	-0.42	0.56	-0.29	0.73	0.61	-0.27	-0.23	11
PIC52021.OUT	12	0.93	-0.61	-0.35	0.61	0.77	0.71	0.63	0.94	0.71	0.06	0.04	1	0.43	0.92	-0.83	-0.92	-0.93	-0.9	-0.85	-0.17	-0.86	-0.32	-0.61	-0.48	-0.53	0.86	-0.53	0.84	-0.52	0.69	-0.07	0.69	-0.67	-0.49	-0.92	12
PIC52022.PV	13	0.52	-0.46	-0.26	0.18	0.53	0.43	0.73	0.54	0.66	0.52	0.48	0.42	1	0.45	-0.2	-0.53	-0.53	-0.59	-0.69	0.02	-0.23	0.02	-0.06	-0.22	-0.51	0.31	-0.18	0.39	-0.18	0.44	0	0.78	-0.13	-0.2	-0.41	13
PI52016.PV	14	0.87	-0.83	-0.38	0.4	0.69	0.68	0.52	0.87	0.76	0.25	0.25	0.82	0.48	1	-0.68	-0.87	-0.88	-0.85	-0.79	-0.08	-0.67	-0.15	-0.42	-0.41	-0.59	0.8	-0.43	0.76	-0.43	0.67	0.02	0.61	-0.48	-0.4	-0.86	14
TI51015.PV	15	-0.8	0.45	0.23	-0.41	-0.69	-0.64	-0.45	-0.61	-0.54	-0.03	0	-0.82	-0.2	-0.98	1	0.8	0.79	0.75	0.67	0.15	0.55	0.8	0.8	0.44	0.64	-0.83	0.47	-0.79	0.49	-0.59	0.07	-0.59	0.64	0.45	0.84	15
TI51021A.PV	16	-0.99	0.81	0.33	-0.54	-0.8	-0.76	-0.64	-0.99	-0.9	-0.38	-0.36	-0.92	-0.53	-0.97	0.8	1	0.99	0.99	0.93	0.17	0.69	0.27	0.51	0.51	0.73	-0.88	0.53	-0.81	0.58	-0.82	0.87	-0.48	0.58	0.5	0.95	16
TI51164.PV	17	-0.98	0.86	0.37	-0.51	-0.8	-0.77	-0.62	-0.99	-0.9	-0.37	-0.34	-0.92	-0.53	-0.98	0.78	0.99	1	0.99	0.92	0.14	0.67	0.24	0.49	0.49	0.71	-0.88	0.53	-0.82	0.54	-0.79	0.04	-0.48	0.57	0.48	0.95	17
TI51156.PV	18	-0.97	0.58	0.28	-0.56	-0.84	-0.74	-0.72	-0.97	-0.91	-0.48	-0.43	-0.9	-0.39	-0.96	0.75	0.99	0.98	1	0.97	0.22	0.72	0.32	0.49	0.52	0.76	-0.84	0.58	-0.77	0.61	-0.85	0.12	-0.58	0.58	0.51	0.92	18
TI51154.PV	19	-0.91	0.52	0.21	-0.6	-0.89	-0.69	-0.83	-0.92	-0.89	-0.49	-0.47	-0.85	-0.69	-0.79	0.67	0.89	0.92	0.97	1	0.27	0.73	0.39	0.45	0.51	0.75	-0.79	0.6	-0.7	0.62	-0.83	0.16	-0.62	0.84	0.5	0.84	19
TI52002.PV	20	-0.18	-0.13	-0.19	-0.49	-0.33	-0.21	-0.35	-0.18	-0.2	-0.23	-0.26	-0.17	0.02	-0.09	0.15	0.17	0.14	0.22	0.27	1	0.77	0.87	0.84	0.86	0.22	-0.14	0.91	-0.11	0.53	-0.39	0.96	-0.09	0.66	0.88	0.15	20
TI52004.PV	21	-0.88	0.27	0.04	-0.71	-0.71	-0.95	-0.62	-0.88	-0.88	-0.37	-0.39	-0.89	-0.23	-0.97	0.55	0.89	0.87	0.72	0.73	0.77	1	0.82	0.7	0.7	0.55	-0.81	0.93	-0.86	0.98	-0.78	0.81	-0.32	0.75	0.72	0.88	21
TI52006A.PV	22	-0.26	-0.17	-0.3	-0.89	-0.38	-0.24	-0.38	-0.25	-0.33	-0.35	-0.39	-0.22	0.02	-0.15	0.2	0.27	0.24	0.32	0.36	0.87	0.82	1	0.54	0.57	0.31	-0.18	0.84	-0.86	0.93	-0.8	0.83	-0.17	0.87	0.8	0.21	22
TI52008A.PV	23	-0.5	0.26	0.18	-0.46	-0.52	-0.5	-0.4	-0.52	-0.23	0.08	0.09	-0.61	-0.08	-0.42	0.6	0.81	0.49	0.49	0.45	0.64	0.7	0.54	1	0.87	0.34	-0.62	0.76	-0.5	0.68	-0.4	0.68	0.08	0.89	0.9	0.54	23
TI52031A.PV	24	-0.48	0.28	0.21	-0.45	-0.57	-0.53	-0.46	-0.5	-0.5	-0.3	-0.5	-0.49	-0.22	-0.41	0.44	0.61	0.49	0.52	0.51	0.65	0.7	0.57	0.87	1	0.55	-0.44	0.76	-0.4	0.7	-0.49	0.73	-0.29	0.89	0.89	0.5	24
TIC52005.PV	25	-0.89	0.4	0.25	-0.46	-0.71	-0.61	-0.56	-0.7	-0.76	-0.68	-0.66	-0.53	-0.51	-0.59	0.64	0.73	0.71	0.76	0.73	0.22	0.55	0.31	0.34	0.56	1	-0.47	0.48	-0.4	0.51	-0.79	0.17	-0.67	0.42	0.53	0.68	25
TIC52021D.OUT	26	0.89	-0.67	-0.45	0.47	0.72	0.79	0.6	0.9	0.85	-0.93	-0.96	0.96	0.31	0.8	-0.83	-0.89	-0.89	-0.84	-0.76	-0.14	-0.81	-0.18	-0.62	-0.44	-0.47	1	-0.49	0.96	-0.48	0.81	-0.38	0.18	-0.67	-0.45	-0.93	26
TIC52021D.PV	27	-0.54	0.19	0.04	-0.82	-0.62	-0.53	-0.55	-0.55	-0.54	-0.33	-0.35	-0.53	-0.19	-0.43	0.47	0.56	0.53	0.58	0.6	0.81	0.85	0.64	0.76	0.76	0.48	-0.49	1	-0.44	0.63	-0.64	0.75	-0.27	0.79	0.79	0.53	27
TIC52025D.OUT	28	0.92	-0.65	-0.47	0.4	0.68	0.84	0.48	0.84	0.95	-0.12	-0.16	0.84	0.89	0.79	-0.79	-0.81	-0.82	-0.77	-0.7	-0.11	-0.56	-0.09	-0.9	-0.4	-0.4	0.96	-0.44	1	-0.39	0.51	0.03	0.09	-0.64	-0.41	-0.89	28
TIC52025D.PV	29	-0.55	0.1	-0.1	-0.89	-0.61	-0.47	-0.56	-0.85	-0.58	-0.38	-0.42	-0.52	-0.15	-0.43	0.46	0.56	0.54	0.61	0.62	0.83	0.86	0.93	0.68	0.7	0.51	-0.48	0.93	-0.39	1	-0.76	0.74	-0.27	0.73	0.72	0.52	29
TIC52027D.OUT	30	0.8	-0.29	-0.03	0.61	0.68	0.55	0.62	0.78	0.61	0.57	0.56	0.69	0.44	0.67	-0.59	-0.82	-0.79	-0.85	-0.83	-0.38	-0.76	-0.6	-0.4	-0.49	-0.79	0.61	-0.64	0.51	-0.76	1	-0.27	0.59	-0.49	-0.48	-0.7	30
TIC52027D.PV	31	-0.06	-0.15	-0.19	-0.4	-0.25	-0.23	-0.3	-0.07	-0.13	-0.19	-0.22	-0.07	0	0.02	0.07	0.07	0.04	0.13	0.18	0.86	0.61	0.82	0.88	0.73	0.17	-0.06	0.79	0.08	0.74	-0.27	1	-0.12	0.89	0.74	0.05	31
TIC52028D.OUT	32	0.44	-0.44	-0.32	0.21	0.49	0.51	0.52	0.48	0.71	0.75	0.73	0.25	0.79	0.41	-0.09	-0.48	-0.48	-0.58	-0.63	-0.08	-0.32	-0.17	0.05	-0.26	-0.47	0.18	-0.27	0.09	-0.27	0.56	-0.12	1	-0.04	-0.25	-0.36	32
TIC52028D.PV	33	-0.58	0.3	0.19	-0.5	-0.6	-0.58	-0.47	-0.6	-0.42	0	0.61	-0.67	-0.13	-0.49	0.64	0.58	0.57	0.58	0.54	0.85	0.75	0.57	0.89	0.88	0.42	-0.67	0.79	-0.64	0.73	-0.48	0.89	-0.04	1	0.91	0.61	33
TIC52030D.PV	34	-0.48	0.25	0.18	-0.44	-0.59	-0.51	-0.44	-0.6	-0.49	-0.27	-0.27	-0.48	-0.2	-0.4	0.45	0.5	0.48	0.51	0.5	0.88	0.72	0.6	0.9	0.89	0.53	-0.45	0.78	-0.41	0.72	-0.48	0.74	-0.22	0.91	1	0.48	34
TIC52034.PV	35	-0.94	0.73	0.5	-0.48	-0.8	-0.84	-0.54	-0.95	-0.81	-0.25	-0.23	-0.92	-0.41	-0.96	0.84	0.96	0.95	0.92	0.84	0.15	0.66	0.21	0.54	0.5	0.68	-0.93	0.53	-0.86	0.52	-0.7	0.95	-0.39	0.81	0.48	1	35
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	